THE IPTC RECOMMENDED MESSAGE FORMAT

1995

1. General
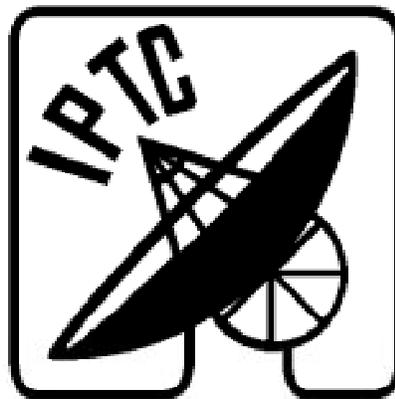
This Recommendation has been formulated by the International Press Telecommunications Council (IPTC) for use in the transmission of messages to newspapers, news agencies and other recipients. It describes the message as it is arriving at the recipient's end. Although designed primarily for computerized information handling, the Recommendation is also suitable for transmission to non-computerized recipients.

The Recommendation has been influenced by the "Highspeed Wire Service Transmission Guidelines" contained in Bulletins 1312 and subsequent amendments thereto of the NAA formerly the American Newspaper Publishers Association (ANPA). Yet because it is intended for international use it takes into account technical and linguistic differences between countries and is designed for use in numerous languages and alphabets.

To provide a degree of flexibility and to minimize changes from earlier practices, some clements in the Recommendation have been designed as "optional" or "recommended". Those not so designated must be complied with when using the Recommendation.

The Recommendation is code transparent. It was originally designed for use in conjunction with versions of the 7-bit coded character set of ISO 646 (CCITT Alphabet Nr. 5), later the possibility for use of alternate character sets (7 and 8 bit), including non-Latin alphabets, has been added (see 7. and Annex C).

IPTC invites you to submit your message format and coded character sets. IPTC will review the submission and issue a Notice of Compliance if the format is technically compatible with this Recommendation. A Notice of Full Compliance will be issued if the interested parties in the respective country do not raise any objections. Submission Forms can be obtained from IPTC. (see Annex D).

2. Format Elements

The format is composed of four sections:

> preheader information,
> message header,
> message text,
> post-text information.

The format elements are separated by the control characters SOH, STX or ETX and terminated by EOT.

3. Pre-Header Information

This is not specified in the Recommendation but may consist of or include:

characters required for transmission purposes, e.g. SYN, selective calling signals to select particular terminals, signals for starting machinery, signals for preparing terminals to receive the message header, e.g. CR and LF, escape sequences for announcing the code extension facilities, the appropriate characters (escape sequences and locking-shift functions) for the initial designation and invocation of the coded character sets used in the following data (see 7. and Annex B and IPTC Register of Codesets).

The composition of the pre-header is at the discretion of the originator except that when SYN characters are used, two of these signals shall immediately precede the SOH in the message header, and if a code extension technique is used. it will follow the ISO 2022 International Standard. SOH, STX, ETX and EOT are not permitted.

4. Message Header

The message header conveys information which is required in the recipient's offices to facilitate handling by editorial and other staff. This information is contained in fields which are separated by field separators which consist of one SP character or CR LF.

4.1. Start of Header

Purpose
To show that header information will follow.

Composition
One SOH character.

Note: A field separator is not required here and must not be used.

4.2. Source Identification and Message Number Field

Purpose
To provide a means to identify the news service of the originator and the message. IPTC recommends that the number of each message be incremented by one from the preceding message in order to enable the recipient to detect the loss of a message.

Composition
Each message transmitted contains a number from 001 to 999 or from 0001 to 9999. These numerals are preceded by one, two or three alpha characters, upper and lower case in any combination, to indicate the source, e.g. "byn379" or "MbsO427". The transmission of leading zeros is at the discretion of the originator.

4.3. Field Separator

One SP character.

4.4. Priority of Story Field

Purpose
To indicate the editorial urgency of a story.

Composition
One numeral from a scale ranging from 1 for the most urgent, 4 for normal to 6 for the least urgent.

4.5. Field Separator

One SP character.

4.6. Category of Story Field

Purpose
To denote the category of a story.

Composition
One, two or three alpha characters. upper and lower case in any combination. e.g. "fin" for financial or "Pol" for political, etc.

4.7. Field Separator

One SP character.

4.8. Word Count Field

Purpose
To show the length of a text.

Composition
The word count is obtained by counting the number of characters after STX and before ETX and dividing by a specified average number of characters per word. This is expressed in up to four numerals. The transmission of leading zeros is at the

discretion of the originator.

4.9. Field Separator

One SP character.

4.10. Optional Information Field

Purpose
To provide supplementary information regarding messages at the discretion of the originator.

Composition
Any characters up to a maximum of 50. CR. LF, SOH, STX, ETX and EOT are not permitted.

4.11.   Field Separator

CR LF   (additional LF characters are permitted at the discretion of the originator).

4.12.   Keyword and Catch-Line Field

Purpose
To index the story contained in the message by a keyword and a catch-line.

Composition
Any characters up to a maximum of 69. It is recommended that. if both. a keyword and a catch-line are used. the keyword precedes the catch-line. CR, LF, SOH, STX, ETX and EOT are not permitted.

4.13.   Field Separator

CR LF   (additional LF characters are permitted at the discretion of the originator).

4.14.   Start of Text

Purpose
To indicate that text will follow.

Composition
One STX character.

5. Message Text

The message text may comprise any characters of the designated coded character set(s). except SOH. STX. ETX and EOT.

Slug information is an optional part of the message text and not subject to this Recommendation, except that:

a)   it is commencing immediately after STX.
b)   it should be terminated by, and therefore should not otherwise contain. a hyphen or colon or equals sign.
c)   it is included in the word count of the message.

Furthermore, it is recommended that sender and recipient of the data agree on the following conventions:

5.1. Mandatory Line Break

to define a mandatory line break i.e. a mandatory line end (e.g. for end of paragraph or end of tabulated line) and/or a new paragraph in one of the following ways in coded character sets which do not use the respective characters for other purposes:

CR or CRLF or LF   if not used for formatting text for a printer

Mandatory line end identifiers:

```
CR LF [CR] LF          O/13  0/10  [0/13] 0/10 (blank line)
< CR LF                3/12 0/13 0/10 (< used as Quad Left)
ETB [CR LF]            1/7 [0/13 0/10] (ETB used as end of paragraph identifier)
```

New paragraph identifiers:

```
CR LF SP SP            0/13 0/10 2/0 2/0 (indent)
[CR LF] HT             [0/13 0/10] 0/9 (HT used as indent)
[CR LF] EM             [0/13 0/10] 1/9 (EM used as EM Space Indent)
CR LF BS               0/13 0/10 0/8 (BS used as Tab Line Indicator TLI)
CR LF ^                0/13 0/10 5/14 (^ used as Upper Rail UR)
CR LF @                0/13 0/10 4/0 (@ used as Lower Rail LR)
```

5.2. No-Break Space (NBSP)

to use the combination GS SP (1/13 2/0) as a No-Break Space in coded character sets which do not provide for a NBSP and which do not use GS for other purposes.

5.3. Soft Hyphen (SHY)

to use the combination RS - (1/14 2/13) as a Soft Hyphen in coded character sets which do not provide for a SHY and which do not use RS for other purposes.

5.4. Open/Close Quotes

5.4.1   If separate codes are used for Open and Close Quotes.

If coded character sets do not provide for Open and Close Quotes (e.g. ISO 646 or ISO 8859-2) to use
           ` (6/0) or `` (6/0 6/0) as Open Quote
          and
          ' (2/7) or ' ' (2/7 2/7) or " (2/2) as Close Quote, or

If coded character sets do provide for Open and Close Quotes to use the codes assigned in that set..

For the ISO 8859-1 or IBM code page 850 to use the "left angle quotation mark" as Open Quote and the "right angle quotation mark" as Close Quote.

5.4.2   If one single code is used for Open and Close Quotes

to use the quotation mark (2/2) for Open and Close Quotes and to apply the following principle:

```
        to treat the combinations
        SP"      (2/0 2/2)
        LF"      (0/10 2/2)
        -"       (2/13 2/2)
```

as Open Quotes, all other combinations with 2/2 as Close Quotes.

5.4.3   Sub Quotation Marks.

If a Subquote is inserted in a Quote using either of the above recommendations for a Quote, the quotation mark (i.e. 2/2 in ISO 8859-1) should be used for opening and closing a subquote, applying the principle of code combinations as in 5.4.2

5.5. EM Dash

to use the character _(5/15) or the combination -- (2/13 2/13) as EM Dash in coded character sets which do not provide for an EM Dash.

5.6. Tabular   Material

to use the character US (1/15) as a data-field (columns) separator in tabular material, e.g.

<US>IBM   <US>98.52   <US>98.52   <US>
<US>Philips   <US>101.03   <US>100.01   <US>-.99

double column:

<US>City   <US><US> Max./ Min.   <US>Forecast
<US>Bruxelles   <US><US> 30/29   <US> Clear
<US>London   <US><US> 10/11   <US> Rain
<US>Zurich   <US><US> -1/-10   <US> Snow

An unlimited number of SP characters may be included to allow visual alignment on teleprinters or other such devices, but a separator character US implies that preceding or succeeding SP characters be discarded by the typesetter.

Any blank field (where the data is missing and not presumed to be of zero value) will consist of a US character and no other information. The separator character must be maintained to ensure format consistency.

If a Central Field Separator CFS (e.g. for side-by-side box scores) is necessary, the character FS (1/1) should be used in coded character sets which do not use 1/12 for other purposes.

6.   Post-Text Information

6.l. End of Text

Purpose
To show the end of the message text.

Composition
CR LF ETX.

6.2. Date and Time

Purpose
To show the date and the time at which the transmission of the text ended.

Composition
A six-numeral word consisting of two numerals each for the day, hour and minute.

Time Zone, Month, Year
Optionally this word may be extended by three upper or lower case alpha characters to indicate the zone of date/time and three further upper or lower case alpha characters to show the month and two numerals for the year. If used each of these elements must be separated by one SP character, e.g. 10.45 p.m. on January 7th, 1991 would be shown either as 072245 or as 072245 jan 91 or as 072245 GMT or as 072245 gmt JAN 91.

6.3. Message Separation Pattern (optional)

Purpose
To separate messages in teleprinter reception.

Composition
Any characters up to a maximum of 32. SOH. STX, ETX and EOT are not permitted.

6.4. End of Transmission
Purpose
To indicate that transmission of the message has been completed.

Composition
One EOT character.

7.   Codes (Coded Character Sets)

If no code extension technique is used and if sender and recipient have not agreed on another Code(s) to be used, the use of ISO 646 IRV (in 7-bit environments) or ISO 4873 DV (in 8-bit environments) is assumed (see below).

If a code extension technique is used or if sender and recipient have agreed on the Code(s) to be used, then a recognized coded character set of the IPTC Code Library shall be used in any case. This library consists of two registers, the *International Register of Coded Character Sets to be used with escape sequences* and the *IPTC Register of Codes*. Both registers can contain a *Repertoire of Graphic sets*, a *Repertoire of Control functions* and a *Repertoire of other coding systems* (e.g. complete Codes). The announcement (if necessary) of the Code extension facilities used in the data which follows and the designation and invocation of Codes is done according to the ISO 2022 International Standard (Code extension techniques, see Annex B) - except with IBM codes!.

The announcement (if necessary) is transmitted in the preheader. Designation and invocation (shifting) can be transmitted anywhere in the message where the respective control characters are permitted. However, it is recommended to transmit, in the preheader, an initial designation and invocation, i.e. to define all designations and the shift status currently in use by transmitting the appropriate escape sequences and locking-shift functions.

IPTC recommends to use either ISO 646 versions or coded character sets of the ISO 8859 series.

	0	1	2	3	4	5	6	7
0	NUL	DLE	SP	0	@	P	`	p
1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3	#	3	C	S	c	s
4	EOT	DC4	¤	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
10	LF	SUB	*	:	J	Z	j	z
11	VT	ESC	+	;	K	[	k	{
12	FF	IS4	,	<	L	\	l		
13	CR	IS3	-	=	M	]	m	}
14	SO	IS2	.	>	N	^	n	~
15	SI	IS1	/	?	O	_	o	DEL

	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	NUL	TC7 (DLE)	SP	0	@	P	`	p								
1	TC1 (SOH)	DC1	!	1	A	Q	a	q								
2	TC2 (STX)	DC2	"	2	B	R	b	r								
3	TC3 (ETX)	DC3	#	3	C	S	c	s								
4	TC4 (EOT)	DC4	¤	4	D	T	d	t								
5	TC5 (ENQ)	TC8 (NAK)	%	5	E	U	e	u								
6	TC6 (ACK)	TC9 (SYN)	&	6	F	V	f	v								
7	BEL	TC10 (ETB)	'	7	G	W	g	w								
8	FE0 (BS)	CAN	(	8	H	X	h	x								
9	FE1 (HT)	EM	)	9	I	Y	i	y								
10	FE2 (LF)	SUB	*	:	J	Z	j	z								
11	FE3 (VT)	ESC	+	;	K	[	k	{								
12	FE4 (FF)	IS4	,	<	L	\	l										
13	FE5 (CR)	IS3	-	=	M	]	m	}								
14		IS2	.	>	N	^	n	~								
15		IS1	/	?	O	_	o	DEL								

NOT USED

ISO 646 IRV	ISO 4873 DV
(International Reference Version)	(Default Version)

1) No character is allocated to positions 00/14 and 00/15

Note:
The difference between ISO 646 IRV and ASCII is only that in the ISO 646 IRV the CURRENCY SIGN (¤) occupies position 2/4 where in ASCII the DOLLAR SIGN ($) is allocated to position 2/4.

LIST OF FORMAT ELEMENTS

PREHEADER INFORMATION (OPTIONAL)

At the discretion of the originator, preheader information may e.g. comprise signals for selective calling, starting machinery, preparing the terminal, transmission purposes, announcing the code extension facilities, the initial designation and invocation of the Code(s) used in the following data. When SYN characters are used, two of these signals shall immediately precede the SOH in the message header.

Name	Example	Composition
MESSAGE HEADER		
Start of Header	SOH	SOH
Source Identification	byn	one, two or three alphas
Message Number	0178	three or four numerals
Field Separator	SP	SP
Priority of Story	2	one numeral (1-6)
Field Separator	SP	SP
Category of Story	pol	one, two or three alphas
Field Separator	SP	SP
Word Count	195	one to four numerals
Field Separator	SP	SP
Optional Information	any characters	up to 50 characters (optional)
Field Separator	CR LF	CR LF (additional LF optional)
Keyword/Catch-Line	any characters	up to 69 characters
Field Separator	CR LF	CR LF (additional LF optional)
Start of Text	STX	STX
MESSAGE TEXT	**Text**	**Text**
POST-TEXT INFORMATION		
End of Text	CR LF ETX	CR LF ETX
Date and Time	071045	six numerals
Field Separator	SP	SP (optional)
Time Zone	GMT	three alphas (optional)
Field Separator	SP	SP (optional)
Month of Transmission	jan	three alphas (optional)
Field Separator	SP	SP (optional)
Year of Transmission	91	two numerals (optional)
Message Separation Pattern	CR LF LF LF	up to 32 characters (optional)
End of Transmission	EOT	EOT

Notes:

1.   When an optional field is not included in a particular transmission, neither a substitute for the field information nor the preceding field separator shall be signalled.

2.   Where the composition of a field is optional, recommended or at the discretion of the originator, it is the responsability of the originator to establish standards acceptable to its subscribers.

3.   Fill characters such as DEL and NUL have no relevance to the format itself.

4.   Check messages and other information of non-editorial content should be transmitted in format. It is recommended to use a particular category indicator for this type of messages.

5.   A single graphic character following two alphas will be considered for use as the Source Identification in Field 4.2.

6.   Where secondary circuit/terminal switching is required, selector characters may be used in Field 4.10. Where selector characters are employed in 4.10., complementary de-selector characters may be used in Field 6.3. but only before any Optional Message Separation Pattern. In neither case may the length of these (4.10. and 6.3.) be extended beyond their prescribed norms.

7.   When a 5-bit character set is used, conversion from 7- or 8-bit Latin alphabet character sets shall be done according to ISO 6936. However, the control characters SOH, STX, ETX and EOT may be substituted by strings such as zczc (SOH), nnnn (EOT) or others. This substitution requires agreement between the sender and the recipient of the data.

IMPLEMENTATION GUIDELINES

Of necessity this Recommendation provides for flexibility. This is because the format has been designed for use in numerous languages and alphabets and takes into account linguistic and national differences.

Members of the press and manufacturers are urged to take advantage of this flexibility as uniformly as possible in the design of their hardware and software.

In future an increasing number of computer systems will be receiving and processing messages on-line from various sources, and the software of such a computer should be based on the agreed guidelines to this format.

By illustration: the program for recognizing the source identity (Section 4.2.) should provide for any combination of 1, 2 or 3 upper or lower case alpha characters. The precise use of each field should be set up at program generation time or by on-line commands.

Furthermore, format elements such as category indicators may be amended by the originator from time to time. The news received in one language may use different mnemonics for indicating a month (6.2.) than in some other language.

Those seeking to adopt the format should always bear in mind that the fields themselves allow for some flexibility as and when necessary but that the field separators themselves are sacrosanct.

ANNEX B

CODE EXTENSION

1. The ISO 2022 International Standard (Code extension techniques)

If other coded character sets than the ISO 646 IRV or ISO 4873 DV (7- and 8-bit environments resp.) are to be used, the Code will be a recognized Code of the *IPTC Code Library*. The Codes will be handied according to the *ISO 2022 International Standard*. This standard defines the technique of *Code extension by increasing the repertoire of characters* in a 3-step operation for graphic sets: 1) a library of Codes is recognized, 2) up to four sets (the G0, G1, G2 and G3 set) are *designated*, 3) the sets are *invoked*, either as a whole or for one character only. For all these purposes the standard defines *escape sequences and shift functions*. The two Control function sets (the C0 and Cl set) are designated and invoked in one step.

The escape sequences used are two-character sequences of the type ESC F or three-character sequences of the type ESC I F (I=intermediate character, always from column 2, F=final character always from column 3 to 7, apart from position 7/15). For further standardization, if more Codes are to be registered than available by these sequences, additional intermediate characters may be added, i.e. sequences of the type ESC I ... I F may be used.

If (a) *Single graphic set(s)* is (are) used only (i.e. up to a maximum of three additional graphic sets and not more than one set to be designated as G0, G1, G2 or G3 set), the designation of set(s) may be agreed between sender and recipient, i.e. no designator escape sequence may be used. Interchanging parties who agree not to use such designators are warned that they may thereby reduce their capability to interchange data subsequently. If *Multiple graphic sets* are used (i.e. more than three additional graphic sets or more than one set to be designated as G0, G1, G2 or G3 set), the designation of the sets is mandatory.

Designation and invocation (shifting) can be transmitted anywhere in the message where the respective control characters are permitted. However, it is recommended to transmit, at the beginning of an information interchange, an *initial designation and invocation,* i.e. to define all designations and the shift status currently in use by transmitting the appropriate escape sequences and locking-shift functions.

At the beginning of an information interchange, in the preheader, it may furthermore be required to *announce the Code extension facilities* used in the data which follows by the appropriate escape sequence (of the type ESC 2/0 F) in order to enable the recipient to interpret the designator escape sequence(s) correctly. It is e.g. possible, if not the full preservation of shift functions is maintained, to define that the escape sequence which designates a set also invokes it.

1.1. Designation and Invocation of Graphic Sets

The Repertoire of Graphic sets contains *Private sequence graphic sets* and *Standardized sequence graphic sets*. They are designated by sequences of the type ESC I F. Private sequence graphic sets have a final character of column 3, standardized sequence graphic sets of column 4 to 7. Intermediate characters 2/8 designate the set as G0 set (94-character sets), 2/9 and 2/13 as G1 set (94- and 96-character sets resp.), 2/10 and 2/14 as G2 set (94- and 96-character sets resp.), 2/11 and 2/15 as G3 set (94- and 96-character sets resp.). The G0 set is invoked by SI, the G1 set by SO (LS1R). the G2 set by LS2 (LS2R), the G3 set by LS3 (LS3R). One single character can be invoked by SS2 or SS3 resp. The coding of LSLR is ESC 7/14, LS2 is ESC 6/14. LS2R is ESC 7/13, LS3 is ESC 6/15, LS3R is ESC 7/12. SS2 is ESC 4/14 (08/14 in 8-bit environments), SS3 is ESC 4/15 (08/15 in 8-bit environments).

The Repertoire of *Multiple-byte graphic sets* contains graphic sets (for Japanese and Chinese e.g., also referred to as Multiple-byte graphic repertoire) designated by sequences of the type ESC 2/4 [I] F and *Dynamically redefinable character sets* designated by sequences of the type ESC I 2/0 F.

**1.2. Designation and Invocation of Control Function Sets and Control Functions**

The Repertoire of Control functions contains *additional C0 sets* of 32 control characters, designated and invoked
by sequences of the type ESC 2/1 F, and *additional C1 sets* of 32 control characters, designated and invoked by sequences of the type ESC 2/2 F. In a 7-bit environment individual control functions of such a C1 set are represented by means of ESC F sequences instead of a single bit combination.The Repertoire of Control functions contains furthermore *Single additional control functions* which are designated *and* invoked by sequences of the type ESC F. Final characters of column 3 are for single additional private control functions, column 4 and 5 for individual control functions of an additional standardized set (e.g. single shift functions like SS2) and column 6 and 7 for additional *standardized* control functions (e.g. locking shift functions like LS2). Single additional

control functions may be designated and invoked also by sequences of the type ESC 2/3 F.

1.3. Designation and Invocation of Other Coding Systems (e.g. complete Codes)

Other coding systems (which are not necessarily character Codes), e.g. Codes containing all characters needed, both control and graphic characters, are designated and invoked by sequences of the type ESC 2/5 F or ESC 2/5 I F. Final characters of column 3 designate and invoke *Private sequence coding systems*, of column 4 to 7 *Standardized sequence coding systems*. The escape sequence ESC 2/5 4/0 is allocated to, and recommended for, use by such other coding systems for returning to the coding system of ISO 2022.

Further details on ISO 2022 should be obtained directly from ISO.
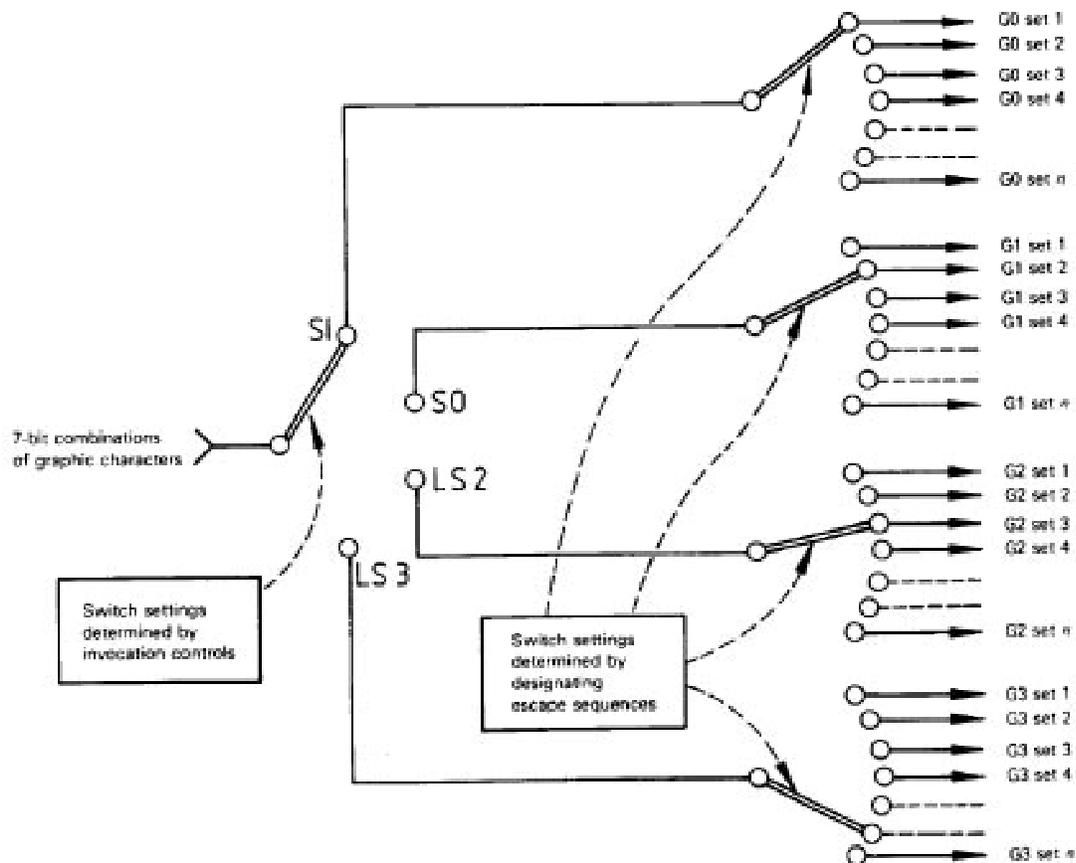
Multiple graphic sets used with locking-shift functions

12
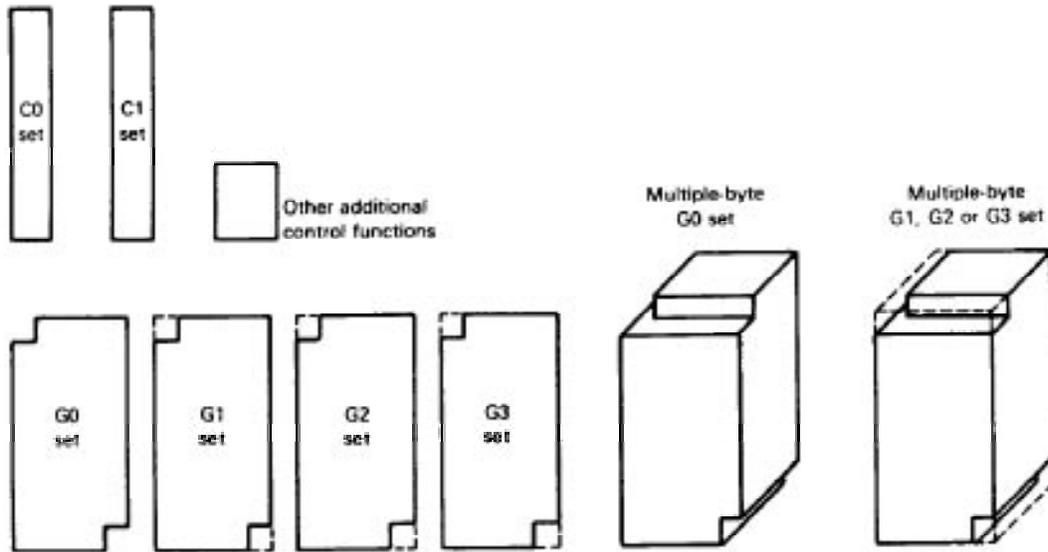
Code extension elements (ISO 2022)

Two-character escape sequences
(ISO 2022)

Three-character escape sequences
(ISO 2022)

Code extension in a 7-bit environment, showing all shift functions (ISO 2022)
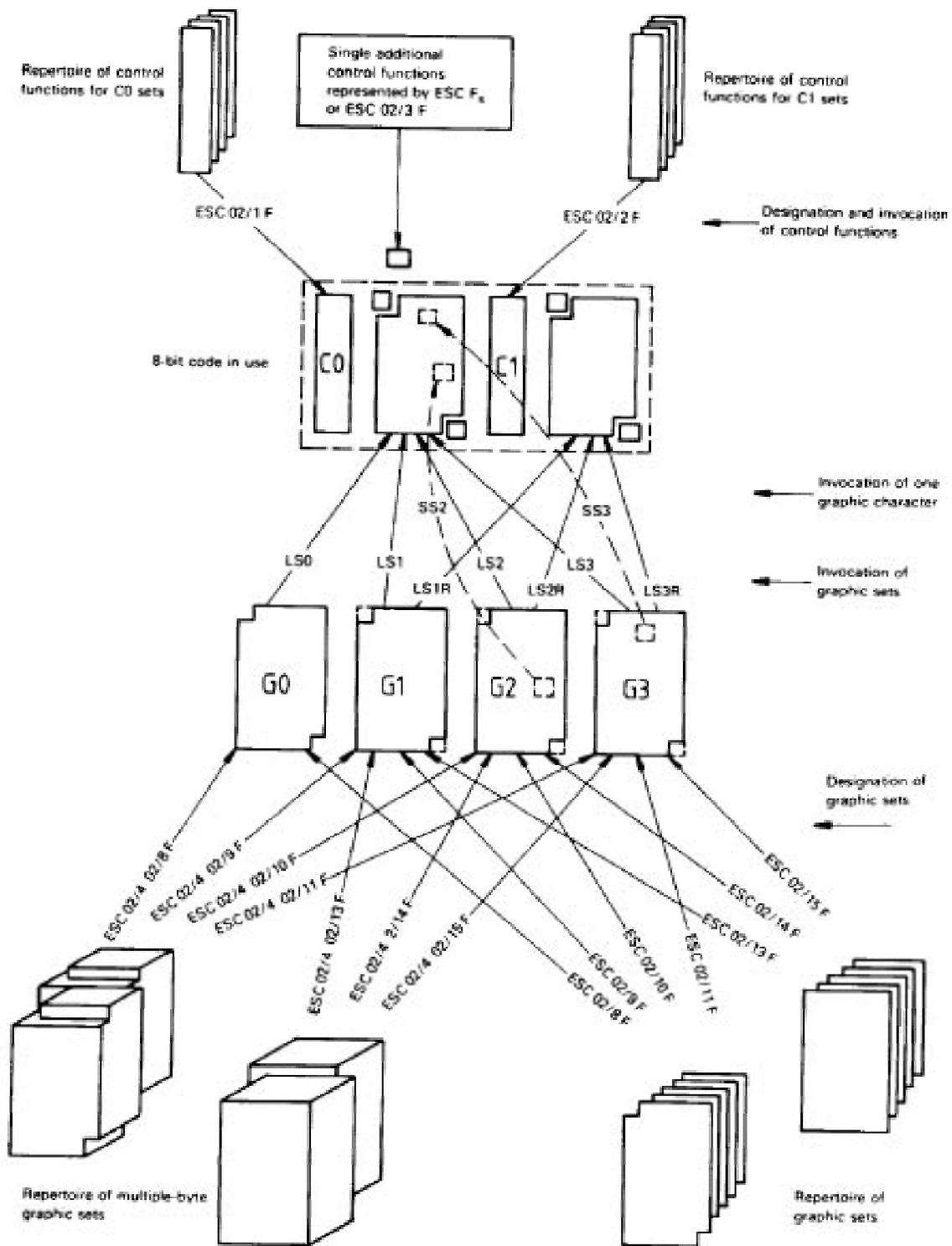
Code extension in an 8-bit environment, showing all shift functions (ISO 2022)

Repertoire of control functions for C0 sets

Single additional control functions represented by ESC F$_s$ or ESC 02/3 F

Repertoire of control functions for C1 sets

ESC 02/1 F

ESC 02/2 F

Designation and invocation of control functions

8-bit code in use

C0

C1

Invocation of one graphic character

SS2

SS3

LS0

LS1

LS2

LS3

Invocation of graphic sets

LS1R

LS2R

LS3R

G0

G1

G2

G3

Designation of graphic sets

ESC 02/4 02/8 F

ESC 02/4 02/9 F

ESC 02/4 02/10 F

ESC 02/4 02/11 F

ESC 02/4 02/13 F

ESC 02/4 2/14 F

ESC 02/4 02/15 F

ESC 02/9 F

ESC 02/10 F

ESC 02/11 F

ESC 02/13 F

ESC 02/14 F

ESC 02/15 F

ESC 02/8 F

Repertoire of multiple-byte graphic sets

Repertoire of graphic sets

15

2. The ISO 646 and ISO 4873 International Standards
(7-and 8-bit coded character sets for information interchange)

The ISO 646 and ISO 4873 International Standards define the technique of Code extension by substitution of certain characters (in 7-bit and 8-bit environments resp.). However, such substitution shall be regarded as constituting a new Code, outside the provisions of ISO 646 and ISO 4873 resp.

Below is the Basic 7-bit Code table of the ISO 646 International Standard (also known as the international Alphabet Nr. 5 as per Recommendation V.3 of CCITT). Following the appropriate rules of the standard, national or application-orientated versions of ISO 646 may be defined. The ISO 4873 International Standard follows exactly the same principles, except that in the ISO 4873 Basic 8-bit Code table no character is allocated to positions 00/14 and 00/15 and that columns 08 and 09 are reserved for the allocation of a set of up to 32 control characters and column 10 to 15 are reserved for the allocation of a set of up to 94 graphic characters.

	0	1	2	3	4	5	6	7
0	NUL	DLE	SP	0	③	P	③	p
1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3	# £	3	C	S	c	s
4	EOT	DC4	¤ $	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
10	LF ①	SUB	*	:	J	Z	j	z
11	VT ①	ESC	+	;	K	③	k	③
12	FF ①	IS4	,	<	L	③	l	③
13	CR ①	IS3	-	=	M	③	m	③
14	SO	IS2	.	>	N	③	n	③
15	SI	IS1	/	?	0	_	o	DEL

16

Notes to the ISO 646 Basic 7-bit Code table

1)   The Format Effectors are defined for applications in which horizontal and vertical movements of the active position are effected separately. If a single control character is required to effect the action of CARRIAGE RETURN in combination with a vertical movement, the format effector for that vertical movement shall be used. For example, if the function "new line" (equivalent to the combination of CARRIAGE RETURN and LINE FEED) is required as a single control character, bit combination 0/10 shall be used to represent it. This substitution requires agreement between the sender and the recipient of the data, and the format effectors (LINE FEED, VERTICAL TABULATION and/or FORM FEED) that are affected shall be identified.

In order to avoid the need for such prior agreement, to facilitate interchange and to avoid conflicts with specifications in other International Standards, the use of format effectors for vertical movements is deprecated. It is strongly recommended to use two control characters, for example CARRIAGE RETURN (CR) and LINE FEED (LF) to obtain the effect of "new line".

2)   Two alternative graphic characters are allocated to each of the bit combinations 2/3 and 2/4. Either the character POUND SIGN (£) or the character NUMBER SIGN (#) shall be allocated to bit combination 2/3 and either the character DOLLAR SIGN ($) or the character CURRENCY SIGN (¤) shall be allocated to bit combination 2/4. Unless otherwise agreed between sender and recipient, the graphic symbols £, $ and ¤ do not designate the currency of a specific country.

3)   **National or application-orientated graphic character allocations.** No specific graphic character is allocated to the ten bit combinations 4/0, 5/11 to 5/14, 6/0, and 7/11 to 7/14. These bit combinations arc available for national or application-orientated use. A unique graphic character shall be allocated to each of these bit combinations, or the bit combination shall be declared unused.

If there is in a country no special demand for specific graphic characters, it is strongly recommended that the characters of the International Reference Version (IRV) be selected and allocated to the same bit combinations as in the IRV.

However, when graphic characters that arc different from the characters used in the IRV are required, they shall have distinct forms and be given distinctive names which are not in conflict with any of the forms or the names of any of the graphic characters in the IRV.

Composite graphic characters: In any version of the 7-bit coded character set specified according to this International Standard, all graphic characters are spacing characters which cause the active position to move forward. However, by using BACKSPACE or CARRIAGE RETURN, it is possible to image two or more graphic characters at the same character position.

For example, SOLIDUS (/) and EQUALS SIGN (=) can be combined to image "not equals". The character LOW LINE (_), that may be used as a free-standing character, can also be associated with other character(s) to represent the graphic rendition "underlined".

Diacritical marks can be allocated to the bit combinations specified in 3) and be available for composing accented letters. For such composition, it is recommended to use a sequence of three characters, the first or last of which is the letter to be accented and the second of which is BACKSPACE. Furthermore, QUOTATION MARK ("), APOSTROPHE (') or COMMA (,) can be associated with a letter by means of BACKSPACE for the composition of an accented letter with a diaeresis, an acute accent or a cedilla, respectively.

Further details on ISO 646 and ISO 4873 should be obtained directly from ISO.