a view of the history of the ISO character code

R. W. Bemer
Honeywell Information Systems
Phoenix, AZ, US

FROM THE EDITOR

This is another documented personal recollection of the development of basic data processing tools. The character code is perhaps the most fundamental of all such tools. We can program without COBOL, but not without coded representations of our most common symbols for calculation and language. Even the Japanese have found it necessary to represent their most elementary alphabet, Kata Kana, in code suitable for computer processing.

ORIGINS

Encoding of character sets has taken many forms -- semaphore, Morse code for telegraphs, and mechanical linkages for typesetting machines. It did not become permanent and interchangeable until the advent of perforated tape and punch cards, representing the continuous and discrete media. Except for looms, punch cards were used first for counting and statistical applications, requiring at first only a 12-character set -- 10 digits and 2 signs. Perforated tape was used first for message communication, and the 5-track tape offered 60 characters -- 30 encodings had double meaning following Figure Shift and Letter Shift, the other two combinations. This code was due to Baudot, and the assignment of codes to characters was made (for alphabet only) on the basis of letter usage frequency, to conserve electrical energy.

Both perforated tape and cards, being available, replaced manual switches as input devices for computers (cards predominating in the US, perforated tape in Europe and elsewhere). Most early computing work was computational, and the punch card set did not receive much impetus for enlargement until application to business problems.

In the late 1950s it became apparent that code expansion had grown uncontrolled. There were more than 20 variations of the punch card code [129,131], and more than 60 different internal representations in computers [12].

The Baudot Code was not without its variations, although the CCITT (Consultative Committee International for Telegraph and Telephone) Working Alphabet No. 2 was used for interchange throughout much of the world. Even so, the binary positional notation was perturbed by different assignments of bits to the tracks.

PRESSURES AND MOTIVATIONS

Some early pressures for code expansion came from programming languages. Univac I, the first computer with alphabetic input, had a 51-character set on its printer. The IBM set had mechanical keypunch limitations to 48 characters, mostly business symbols. Starting in 1954, the need for FORTRAN (mathematical and arithmetic) symbols to take the same encodings as these business symbols led to much confusion in installations that did both types of work. Pressures arose for unique encodings [1,3,4,126]. In 1956 September it was decided that the IBM Stretch computer would have a 64-bit word,and that the characters would be represented by 8 bits [19,70]. This decision had much influence on subsequent computer architecture. The Los Alamos group devised its own input character set, and IBM built a special entry typewriter for it [10].

The advent of ALGOL in 1958, with (in theory) 116 single symbols, was the impetus for many proposals. H. S. Bright proposed one such set to the IBM SHARE organization on 58 September 11, as a "point of departure for hardware implementation of the IAL" [4]. Often one to be ahead of his time, his tableau showed the European "one" and the letter "oh" with a curlicue. Yet he objected to SHARE asking IBM for more than 128 characters presentable in media and equipment unless someone was able to prove the need. Others foresaw the need very well, and there was much work going on within IBM at the time, following on the Stretch decision. The 1403 chain printer was being developed, the 240 positions around the chain permitting 2 sets of 120 or 4 sets of 60, in addition to the planned 5 sets of 48 for the traditional IBM set.

vol 6 #4
1972

EARLY STANDARDIZATION EFFORTS

The real future for computers became clearer in 1958 and 1959. Data processing centers were outgrowing the period when information processing was primarily a local affair, and some connection to communications networks had been made. First to emerge was data interchange, as for Social Security Administration and Internal Revenue Service records. Then came program interchange, more difficult because both symbols and encodings varied.

Several independent standardization projects began at about this time:

- The British Standards Institution (BSI), starting with paper tape and punched cards, gradually moving into the general problem.
- SHARE, seeking to coordinate its IBM equipment.
- The Electronic Industries Association, which also started first from the paper tape aspect, partly due to impetus from Numerical Control.
- The US Department of Defense, for whom the Signal Corps developed the Fieldata Code (which, despite drawbacks, was a great improvement on existing codes, and many of its features are seen in the ISO code). The Fieldata Code still exists as the internal code in the Univac 1107, 1108, and 1110.

IBM was also doing much work in codes, for in 1959 there were 9 different internal codes existing in IBM equipment, and interchange on media was a problem. Development of the 8000 series (precursor of the 360) was under way, and my group (in IBM) was planning the coordination between codes and media representation, following on the Stretch work. A 1959 November memo to V.P. John McPherson outlined a plan for a "single character set to be used by all future computer systems ... the code to be the external storage code as well as internal machine code" (i.e., direct representation in punch cards, as well). It was an 8-bit set.

John Gosden saw this work in the Fall of 1959 and arranged for me to be invited to make a presentation to the BSI codes committee, which occurred in 1960 February. This established a very important contact point with Hugh McGregor Ross (for a while the code was known colloquially in ECMA (European Computer Manufacturers Association) circles as the Bemer-Ross code).

EXPANDING THE JURISDICTION

It was natural that IBM, with its rapid multinational growth, would take a strong interest in the standardization process. Jim Birkenstock, V.P. of Commercial Development, drew up plans to revitalize the Office Equipment Manufacturers Institute and convert it to the Business Equipment Manufacturers Association. A vital consideration in this move was the preparation for BEMA to act as sponsor for the US standards effort in data processing for ASA (American Standards Association, now ANSI). To that end, scopes and programs of work for the various activities were drafted at IBM and presented at the 1960 January 13 meeting convened by ASA.

In the ISO (International Standardization Organization) area, it was Sweden (probably Olle Sturen) that recommended, in late 1959, activities in data processing standards. This led to a Round Table Conference in Geneva on 1961 May 16, which in turn led to the formation of ISO/TC97, Computers and Information Processing, for which the US was assigned the secretariat. Several Working Groups were established, in particular WG B, on Coded Character Sets. ECMA had just been formed in 1961 May, and had liaison representation at the Geneva meeting. Its corresponding body for codes was Technical Committee 1.

US CONTRIBUTIONS

The 1960 January meeting of ASA had led to the formation of Sectional Committee X3, with the same name as TC97. X3 held its first meeting on 1960 August 4, and X3.2 (Codes and Input-Output) its first on October 4, chaired by I.C.Liggett of IBM, who was directing IBM efforts (other work continued in the US until X3.2 jurisdiction was established authoritatively). Membership was from the Federal Government, data processing, and communications interests.

The latter were represented primarily by the Bell System, which was to convert its TWX service from manual to dial operation in 1962. Along with this was a new transmission arrangement that enabled 100 word-per-minute service in addition to the existing 60 wpm service. This meant that the majority of teletypewriters (Model 15, limited to 75 wpm) would have to be replaced. To Bell, and to John Auwaerter of Teletype, this seemed an excellent opportunity to escape from the limitations of 5-level Baudot Code and the 3-row keyboard to a 4-row keyboard more similar to conventional typewriters. The Bell System was not only an active participant, but actually hastened development by adopting conclusions as fast as they were reached and incorporating them in their new designs. It even chose the perforated tape track assignment in advance of that standard, and thus made it rather easy to reach agreement because of the volume of Model 33 and 35 equipment in actual operation.

X3.2 did extensive work on development of criteria for a code. It considered aspects such as collating sequence and natural orderings of subsets, keyboard design considerations, and programming effects such as manipulation of graphics by classes, simpler language scans, and more graphics for clarity of printed output. This was all based upon the assumption of a dense 6-bit code, as used by almost all computers of the time (Stretch was the exception, at 8 bits).

No existing code was found to be satisfactory. The Fieldata Code, as Military Standard 188 [72], was in final stages of adoption in 1959 December, and was offered for consideration. However, it was primarily for communications, although computer equipment was used in the process. For data processing, it had the defect that control codes and graphics were intermingled in 3 of the 4 columns. The alphabet was pushed to the bottom of its columns, leaving no room for additional characters of other alphabets, such as the Scandinavian. The IBM codes had a peculiar and overriding problem - none of them had the same sequence as the standard IBM collating sequence.

THE 7-BIT ERA BEGINS

After trying many months without success to select 64 characters that would satisfy the requirements of both data processing and communications, X3.2 finally settled upon a 7-bit set at its 61 June 7-9 meeting. The alphabet was pushed to the bottom of its second column, as inherited from Fieldata. Nevertheless, the US Department of Defense was beginning to get concerned about conflict with Fieldata. Secretary Macon, in his summary report of the first 9 meetings, suggested that the code be called Fieldata No. 2, to give that credit.

IBM continued to be concerned about aspects of international compatibility. Two sets had been presented at the June 9 meeting with the Department of Defense - one with the alphabet at the bottom, and one with the alphabet starting in the second position, as now (column 1 of Table 1). This second condition was furnished by Ross, for BSI.

At its next meeting X3.2 had to decide between these two proposals. Study of Scandinavian phone books showed the need to have at least 3 extra positions following the 26-letter alphabet. I took this position strongly in the ensuing argument (which, although vociferous, is not indicated in the minutes). Placement of the special symbols was another matter; the traditional IBM collating sequence was specials (S), alphabet (A), digits (D). I had a ready solution - the major special symbols (particularly "minus") were to be taken out of the digit column and placed in the previous column. For ordering purposes, the digit column could then be made to collate higher than the alphabet (for IBM's needs) by passive logic of bit inversion. The specials following the digits had not appeared previously in key fields used by IBM customers, and it did not matter if they were high to the alphabet. The 4-bit subset was thus constructible with a 4-position jog, or displacement. This was also important to IBM, for the packed decimal notation in the new line. On this basis, John Auwaerter and I prepared a presentation that convinced X3.2 to settle the code in its present structure; the basic 7-bit standard was agreed - the alphabet starting in the second position! The draft then stood as in column 2 of Table 1.

I had called a joint meeting of IBM, SHARE, and GUIDE, to regularize the IBM 6-bit set to become the standard BCD Interchange Code [76]. Frequency studies of symbol occurrence had been prepared, particularly from ALGOL programs. The meeting of July 6 produced general agreement on a basic 60-64-character set, which included the two square brackets and the reverse slant, which was chosen to be used in conjunction with "/" to yield 2-character representations for the AND and OR of early ALGOL. This is reflected in the set I proposed to X3.2 on September 18 (column 3). The lower case alphabet was also shown, but for some time this was resisted, lest the communications people need more than the two columns then allotted for control functions.

At the 61 November 8-10 meeting, X3.2 constructed the first formal proposal, X3.2/1 [73] (column 4), with the explanations for its structure and content, including the compatibility with Europe, where the UK standard was at the printers. The arrows were moved up so as to be expendable when replaced by 10 and 11 for pence of the English monetary system of that time.

GOING TRULY INTERNATIONAL

The Americans, less accustomed than now to international standardization work, had their eyes opened at the first ISO/TC97 meeting. At the 1961 Fall Joint Computer Conference I broached the possibility of direct discussions at working level to Philippe Dreyfus of ECMA. This was discussed at the concurrent meeting of X3.2 (December 11-13). Dr. J. Barker of BEMA, as chairman pro tem of X3, was irate with such a lack of national loyalty. Nevertheless, (now) X3.2 Chairman Macon announced that arrangements had been made for Bloom (NCR) and Auwaerter (Teletype) to go to Europe for the period of 62 January 2-22. The resulting frank discussions with BSI and ECMA had very good effect. On January 29, X3.2 received from Hugh McGregor Ross a proposed a compromise that could meet the requirements of ASA, ECMA, BSI and the EEA (Electrical Engineering Association of the UK). Thus the groundwork was laid for the first meeting of TC97 Working Group B on 62 May 4.

H. Feissel of BULL was elected chairman. Present were delegations from France, Germany, Italy, the UK, and the US - plus ECMA and UIC (Union Internationale des Chemins de Fer). Despite the official UK position for its own standard, Ross put forth his compatible proposal. Although Germany wanted the alphabet moved to the first position, there was considerable agreement, and a 6- and 7-bit code were resolved. The latter was in two versions, A and B (columns 5 and 6 of Table 1), differing only in the two control columns. Germany preferred A, all others preferred B.

The great significance of moving the work to the ISO level is evidenced in these proposals, noting how many choices remain fixed after this time. Furthermore, the controls were regularized and grouped to 7 transmission controls, 6 format effectors, and 5 device controls; the improvement from the haphazardness of the previous proposals is quite apparent.

It was noted subsequently that the positions of ":" and ";" were unsatisfactory for pence replacement in sterling. Cooperation was so well established that interchange with "*" and "+" was effected by a telephone call to ECMA in Europe, who confirmed its agreement to WG B on June 27.

X3.2, whose written vote on X3.2/1 was all affirmative except IBM and an abstention from ITT, proposed to X3 that the new document, X3.2/4 (column 7), be supported, as these were minor changes. International consistency with Proposal B was stressed, and X3.2 recommended extending the X3 ballot period to September 15, as WG B needed official opinions for the October plenary meeting of TC97.

There followed a difficult period in the US, eventually leading to the first American Standard Code, X3.4-1963.

GETTING AN AMERICAN STANDARD CODE

The original negative vote of IBM on X3.2/1 (at the X3.2 level) was cast by Robert Blue, who had replaced H. J. Smith, Jr., as the IBM representative. Smith was asked, however, to remain as a consultant to X3.2 in recognition of his many fine contributions. This was permitted by IBM, with the explanation that "We have tried to participate ... according to ASA rules, and have allowed our participants to act freely on a purely technical and professional basis. Since the actions of our participants have been construed as company actions and endorsements, we shall, in the future, insure that these are one and the same .." (Bartelt to Macon on March 2). Accordingly, Blue's negative stated that "IBM studies have shown that the proposed 7-bit code with its contiguous alphabet would be expensive to implement in the present domestic data processing environment oriented to the use of 6-bit processors and punched cards". IBM work on 8-bit codes or the new product line was not mentioned.

William E. Andrus, Jr., IBM Group Director of Standards, explained IBM's negative vote in X3: "After extensive study of the entire problem of code standardization, this investigation, together with the experience gained during our many years in the data processing field, indicates that the benefits derived from the adoption of this standard would not justify the economic burdens involved with its implementation. We submit that the BCD Interchange Code [76] is, and must be formally recognized as, a standard because of its wide use both in punch cards and magnetic tape ... IBM supports the development of standard codes (sic plural) for information interchange, but will not support any proposal for a single standard code. X3.2 should consider the possible interaction of future 8-bit codes on present standardization activities".

ITT Communications Systems also took a negative position. A 62 June 18 paper by L. A. DeRosa said that "The proposed code standard ASCII is considered unsuitable as a standard" (it wasn't Fieldata).

As decisions moved from the technical realm to the administrative levels of X3 and the BEMA Data Processing Group (DPG), support lessened. Friden said it must vote negative because terminal costs would be increased 30% to support ASCII (but presented no data). Many minority position papers were written, and X3.2 made extensive and detailed replies to the (some less than rational) objections. The paper of Blue (IBM) and Ungar (ITT), for example, observed that "positive numbers would collate lower than negative numbers". X3.2 took 2 pages to prove that, no matter what positions were assigned to + and -, it is still impossible to order numerals with a binary sort!

Another objection was that it would cost $124 million and take 2 years to change just the 465L computers from Fieldata, a position not supported by the owner of those computers - the US Department of Defense. Thomas D. Morris, US Asst. Secretary of Defense (Installations and Logistics) wrote to X3 on 62 August 6, saying that "This office is wholeheartedly in support of the ASA Sectional Committee X3 standards development effort both domestically and internationally". He recognized the differences with, but similarity to, Fieldata in organization and structure.

On 62 July 29, Lin Griffin, now chairing X3.2, informed the Chairman of X3, C. Phillips, that X3.2 had approved (by majority vote) a motion that none of the minority positions nor the negative and abstaining votes were deemed necessary or sufficient reasons to warrant revision of the proposed ASCII [75]. "The subcommittee reaffirms its belief that the proposed standard code is the best code it can develop for information interchange purposes, and has the best chance of acceptance and implementation on a national and international basis".

On August 17 IBM abandoned protection of the 6-bit code and exposed a paper by C. Mackenzie, "The construction of 8-bit coded character sets", to the BEMA DPG, and to X3.2 on August 22. Here it was admitted that although the IBM Standard BCD Interchange Code did *not* have binary collatability embodied, this proposal did. In short, IBM furnished the prototype of the EBCDIC (Extended BCD Interchange Code). X3.2 rejected it by a 7-1-1 (yes-no-abstain) vote. "Since it is based upon a structure which was rejected earlier by X3.2, the subcommittee does not recommend revision or withdrawal of X3.2/4".

IBM then asked for an evaluation of economic impact, which was rejected as having no precedent in the records of ASA. At the September 10 meeting of X3, postponed by Phillips from August 21 because of what he felt was a current lack of consensus, X3 voted to send Document X3.2/13 (containing editorial revisions to meet requirements of X3.3 for data communication, thus changing the AIEE vote to affirmative) to ASA for approval. The vote was 20-4-5, with the incognito vote of the BEMA members allocated 6-2.

At this point one might expect that the standard would go to ASA for processing and publication, but there were several more hurdles. First was the BEMA DPG, sponsor of the X3 work. Andrus of IBM wrote this group on October 11 to say "We find it necessary to continue to abstain from voting on the X3.2 code until our request of August 17" (for the IBM code to be named a de facto standard) "is acted upon ... be assured that if BEMA will not act upon our requests, and if the X3.2 Code is issued as an American standard by ASA, IBM will cooperate, making the code available where the economic disadvantages to IBM customers are not a factor".

On October 31 the BEMA Standards Review Board voted 11-3-3 for forwarding. Nays were from Friden, Addressograph-Multigraph, and Monroe, all of whom were represented on both X3 and X4 (X4, Office Machines, chaired by MacFarlane of IBM, had objected to the code as it might apply to office equipment). Abstentions were from IBM, Moore Business Forms, and Smith-Corona-Marchant. The RCA position was unknown, but computer manufacturers voting yes included Honeywell, Burroughs, NCR, and Univac.

On 62 November 5, Frank White of the Air Transport Association wrote to ask: "Rumor has it that BEMA has turned down the recommended action of X3 relative to the standard code for data processing. Do you plan to provide ... details ... so that we may take under advisement what may be done to achieve effective implementation of the policy we recommended?" Phillips replied that "The BEMA Standards Review Board decided to defer the transmittal of the proposed Standard Code for Information Interchange until something further was done toward achieving a consensus. The Board also recommended that X3.2 take appropriate action to submit de facto standard codes concurrently with the pASCII (*p*roposed ASCII). It is believed that such actions would satisfy the objections raised by the X4 Sectional Committee". He further notified the Joint Steering Committee of X3, X4, and X6 that the "Sponsor requests that X3 and X4 propose a de facto standard or standards for consideration concurrently with proposals for pASCII expressed in media".

col.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
document	X3.2 to DoD	X3.2	RW8	X3.2/1	TC97 WG 8 A	TC97 WG B B	X3.2/4	ISO #1	STD X3.4 -1963	ISO #2	ISO #3	X3.2.4 /64	CCITT	X3.2 /172	X3.2 /206	ISO #4	ISO DR 1052	X3.2 /384	X3.2 /411
date	1961 06 29	1961 09 14	1961 09 18	1961 11 10	1962 05 04	1962 05 04	1962 05 25	1963 01	1963 06 17	1963 12	1964 05	1964 10 20	1964 10 09	1964 12 03	1965 01 21	1965 03	1966 04	1966 05 20	1966 07 15
0/0		NULL													NUL			NUL	NUL
0/1		RU	VT	SOM	EOT	SOM					SOH								
0/2			rcv 1 on	m off	EOM	EOA					STX								
0/3			rcv 1 off	RU	EOA	EOM					ETX								
0/4		WRU			SOM	EOT													
0/5		motor 1 stop	tmt stop	EOT	WRU						ENQ	ENQ	WRU	ENQ					
0/6		VT	feed out	HT	RU						ACK								
0/7		HT	HT	FF	BELL										BEL			BEL	BEL
0/8		feed out	motor stop			FE_0				BS									
0/9		tmt start	rcv 2 on		FF	HT	(sk)		(sk)										
0/10		BELL			VT	NL	LF												
0/11		rec 1 on	rcv 2 off		LF	VT											Assignments		
0/12		rec 1 off	tmt start	CR	HT	FF											Permanent		
0/13		rec 2 off	RU	LF	CR														
0/14		rec 2 on		UC	SO														
0/15		tmt stop		LC	SI														
1/0				R1 on		DC_0	DC_0	DLE	DC_0	DLE									
1/1				R2 on	DC_4 stop	DC_1													
1/2				X on	DC_3	DC_2													
1/3		data	data	EOA	DC_2	DC_3													
1/4		delimiters	delimiters	R1 off	DC_1	DC_4 stop						DC_4		DC_4					
1/5				R2 off	ERROR	ERROR	ERR	ERROR	ERR	ERROR	NACK				NAK			NAK	NAK
1/6				X off	SYNC										SYN			SYN	SYN
1/7				EOM	EOB	EOB	LEM				ETB								
1/8					S_5 LEM	S_0	S_0	IS_0	S_0	IS_0	CNCL				CAN			CAN	CAN
1/9					S_4						EM								
1/10				data	S_3						SS						SUB		
1/11				delimiters	S_2						ESC								
1/12		LF	LF		S_1						IS_4 FS								
1/13		CR	CR		S_0						IS_3 GS								
1/14		UC	UC								IS_2 RS								
1/15		LC	LC			S_7	S_7	IS_7	S_7	IS_7	IS_1 US								

Steps in the Evolution of the ISO Code

col.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
document	X3.2 to DoD	X3.2	RWB	X3.2/1	TC97 WG B A	TC97 WG B B	X3.2/4	ISO #1	STD X3.4 -1963	ISO #2	ISO #3	X3.2.4 /64	CCITT	X3.2 /172	X3.2 /206	ISO #4	ISO DR 1052	X3.2 /384	X3.2 /411
date	1961 06 29	1961 09 14	1961 09 18	1961 11 10	1962 05 04	1962 05 04	1962 05 25	1963 01	1963 06 17	1963 12	1964 05	1964 10 20	1964 10 09	1964 12 03	1965 01 21	1965 03	1966 04	1966 05 20	1966 07 15
2/0	b̸	→	→	→	SP	SP	b̸	SP	b̸	SP									
2/1	!	'	→	→	!														
2/2	"																		
2/3	?	<	→	→	#														
2/4	$	>	→	→	$	→	→	CS	$	CS	CS_1	$	¤	$	$	CS_1	$		
2/5	'	$	→	→	%														
2/6	&	%	→	→	&														
2/7	—	&	→	→	'	→	→	→	→	'									
2/8	(																		
2/9	)																		
2/10	;	:	:	.	:	:	*										Assignments Permanent		
2/11	:	*	*	,	;	;	+												
2/12	*	/	/	*	,														
2/13	/	-	-	:	-														
2/14	,	→	→	-	.														
2/15	.	→	→	/	/														
3/10	+	;	;	↑	*	*	:												
3/11	⌐	+	+	←	+	+	;												
3/12	<	?	?	+	<														
3/13	>	=	=	;	=														
3/14	X	←	←	=	>														
3/15	[	↑	↑	?															
4/0	=	!	!	!	@	→	→	→	→	→	—	—	`	→	→	→	@		
5/11	≠	#	#	@	[														
5/12	#	@	@	\															
5/13	%	≠	[	[	]						CS_2		~	→	→				
5/14	↑	≤	]	]	↑	→	→	→	→	^									
5/15	←	≥	\	#	←	→	→	→	→	\	→	→	—						
6/0																			
7/11		ansr back	ansr back							—	@	→	→	→	→	→	`		
7/12				ACK			ACK		ACK			{		{			{		
7/13												|		¬			|	|	¦
7/14	ESC	→	→	→	→	→	→	→	→	→		}		}			}		
7/15	DEL											¬		|			—	¬	~

International Reference Version (in OCR-B font)

b_4	b_3	b_2	b_1									
				b_7	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1
					0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	TC$_7$ (DLE)	SP	0	@	P	`	p
0	0	0	1	1	TC$_1$ (SOH)	DC$_1$	!	1	A	Q	a	q
0	0	1	0	2	TC$_2$ (STX)	DC$_2$	"	2	B	R	b	r
0	0	1	1	3	TC$_3$ (ETX)	DC$_3$	#	3	C	S	c	s
0	1	0	0	4	TC$_4$ (EOT)	DC$_4$	¤	4	D	T	d	t
0	1	0	1	5	TC$_5$ (ENQ)	TC$_8$ (NAK)	%	5	E	U	e	u
0	1	1	0	6	TC$_6$ (ACK)	TC$_9$ (SYN)	&	6	F	V	f	v
0	1	1	1	7	BEL	TC$_{10}$ (ETB)	'	7	G	W	g	w
1	0	0	0	8	FE$_0$ (BS)	CAN	(	8	H	X	h	x
1	0	0	1	9	FE$_1$ (HT)	EM	)	9	I	Y	i	y
1	0	1	0	10	FE$_2$ (LF)	SUB	*	:	J	Z	j	z
1	0	1	1	11	FE$_3$ (VT)	ESC	+	;	K	[	k	{
1	1	0	0	12	FE$_4$ (FF)	IS$_4$ (FS)	,	<	L	\	l	|
1	1	0	1	13	FE$_5$ (CR)	IS$_3$ (GS)	−	=	M	]	m	}
1	1	1	0	14	SO	IS$_2$ (RS)	.	>	N	^	n	‾
1	1	1	1	15	SI	IS$_1$ (US)	/	?	O	_	o	DEL

Official Reference Version — R646 (in OCR-B font)

b_4	b_3	b_2	b_1									
				b_7	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1
					0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	TC$_7$ (DLE)	SP	0	③	P	` ④	p
0	0	0	1	1	TC$_1$ (SOH)	DC$_1$	!	1	A	Q	a	q
0	0	1	0	2	TC$_2$ (STX)	DC$_2$	" ⑥	2	B	R	b	r
0	0	1	1	3	TC$_3$ (ETX)	DC$_3$	£(#) ②	3	C	S	c	s
0	1	0	0	4	TC$_4$ (EOT)	DC$_4$	$(¤) ②	4	D	T	d	t
0	1	0	1	5	TC$_5$ (ENQ)	TC$_8$ (NAK)	%	5	E	U	e	u
0	1	1	0	6	TC$_6$ (ACK)	TC$_9$ (SYN)	&	6	F	V	f	v
0	1	1	1	7	BEL	TC$_{10}$ (ETB)	' ⑥	7	G	W	g	w
1	0	0	0	8	FE$_0$ (BS)	CAN	(	8	H	X	h	x
1	0	0	1	9	FE$_1$ (HT)	EM	)	9	I	Y	i	y
1	0	1	0	10	FE$_2$ (LF) ①	SUB	*	:	J	Z	j	z
1	0	1	1	11	FE$_3$ (VT) ①	ESC	+	;	K	③	k	③
1	1	0	0	12	FE$_4$ (FF) ①	IS$_4$ (FS)	, ⑥	<	L	③	l	③
1	1	0	1	13	FE$_5$ (CR) ①	IS$_3$ (GS)	−	=	M	③	m	③
1	1	1	0	14	SO	IS$_2$ (RS)	.	>	N	^ ④⑤	n	‾ ④⑤
1	1	1	1	15	SI	IS$_1$ (US)	/	?	O	_	o	DEL

	ISO	ECMA	ANSI	US FIPS PUB	JIS	USSR	UK
Binary code for characters (plus control meanings)	R646 (67 Dec)	ECMA-6 3rd Ed. (70 Jul)	X3.4-1968	1	C6220-1969	GOST 13052-67	
Hollerith card code	R2021 (71 May)	ECMA-20 (69 Jun)	X3.26-1969	14			BS4636/3 1971
Graphics for the controls	DR2047	ECMA-17 (68 Nov)	X3L2/987 (70 Jul 29)				
National usage	97/2/525		X3L2/1238A				
Additional controls			X3L2/1198				
Serial transmission	R1177 (70 Jan)		X3.15-1966				
Track assignment on 25.40 mm (1") punched tape	R1113 (69 Sep)	ECMA-10 2nd Ed. (70 Jul)	X3.6-1965	2	C6221-1969		BS3880/3 1971
Track assignment on 12.70 mm (0.5") magnetic tape	R962 (69 Feb) R1863 (71 May)	ECMA-12 2nd Ed. (70 Jun)	X3.22-1967	3	C6222-1969		

SOURCE DOCUMENTS - STANDARDS AND DRAFT STANDARDS

Phillips then wrote to Roger Gay of the ASA on December 5, saying "The problem here is one of attempting to introduce a new and nonexistent code convention as the first and only standard in the character code field of information processing systems. Such an action would render widespread de facto code practices "nonstandard" and obsolete, with potential economic and technological hardships to at least the suppliers of equipment affected by such standardization". X3 was informed of the sponsor's actions as of this date.

X3.2 was as distressed as X3. At its 62 December 5-7 meeting, a motion passed 5-2-3 that "X3.2 is of the opinion that if any existing coded character set is to be considered a de facto standard, many must be considered, and that, if any one is found allowable jointly with the pASCII, many will be found allowable ... Analysis of documents [12,129, 131] supports this position". Griffin then prepared a summary memo for the Joint Steering Committee, dated December 18.

Gay of ASA replied to Phillips on 62 December 21: "It is my understanding that ... each of the existing codes had serious shortcomings which precluded its being recommended ... the Constitution of the ASA specifically prohibits the promulgation of conflicting standards ... In X3, however, there was never a formal recognition to any of the 'de facto' codes, and the very insertion of such codes in the American Standard may promote their continued use rather than urge their discontinuance. In other words, you would be maintaining the status quo which has been greatly criticized in the past". He noted that James Birkenstock, V.P. of IBM, had said that "IBM believed industry as a whole stood ready to cooperate in coming up with a single language", and William Suchors, V.P. of Remington Rand "too felt that all manufacturers could harmonize their thinking in the new project". He then reminded BEMA that "At all levels, the views of the sponsor are taken into account. However, the sponsor has no veto power". He advised calling a meeting, and that "if the decision of X3 is that de facto codes should *not* be added, BEMA should transmit the pASCII without delay to ASA".

The Underwood Corporation accented the anomaly of the situation by writing to ASA to say that they were against inclusion of de facto standards but, if there were to be any, here was theirs! Others took frontal action. Paul Riley, Deputy Asst. Secretary of Defense, in his 63 January 18 letter to Phillips, urged approval of ASCII, particularly to put the US in a better position at ISO. G. D. Osborn, of the Executive Office of the President, Budget, wrote X3 members saying that it was important that the proposed code be forwarded to ASA and adopted now.

At the 63 January 24 meeting of X3 it was voted, by a 2 to 1 majority, to transmit the pASCII directly to ASA without recommendation of the sponsor, the first (and perhaps only) time such an action had occurred in the history of the American Standards Association!

In a final attempt, Phillips wrote to Gay on February 8, saying "It appears that a part of the action ... of October 31 was in conflict with this interpretation of ASA regulations ... the sponsor's new recommendation on the pASCII was developed as ... 'The sponsor accepts the principle of a single code for information interchange. This code will be universally useful only when it can be adapted to the common methods of machine-to-machine communication, i.e., paper tape, punched cards, and magnetic tape. The sponsor, therefore, believes that submission of the pASCII to ASA is premature' ". This position, adopted by a 12-3 vote of the BEMA Standards Review Board, was ineffective, however. X3.4-1963 [79](column 9) was approved by ASA on June 17, after a hearing by its Miscellaneous Standards Board.

The action was newsworthy. Ted Merrill of Business Week Magazine made a significant point [21] when he said that "ironing the kinks out of pASCII took about $3 million worth of man-hours". He noted that AT&T was one of the earliest to gamble, with its new Teletype line [29], and that "some equipment manufacturers, notably IBM, agree the basic code is a good one, but would like to wait and see how it will be standardized for punched cards before they give their stamp of total approval". He suggested that "many computers will undoubtedly use it as their own internal language to avoid the translation problem altogether". Adoption of the ASCII was the occasion for a number of other articles publicizing it [24,25,26].

THE ERA OF COOPERATION AND GOODWILL

A special issue of IBM's Data Processor [22] was devoted to describing the adoption of ASCII and its import. "To support ASCII ... IBM plans to provide whatever means are practical to meet customer needs for using the standard code as soon as possible, after media standardization is approved by the American Standards Association. Because standardization is so important, we have already begun to implement ASCII ... We as a corporation are determined to move ahead with ASCII at the most rapid pace possible so that our customers can gain the benefits of standardization across the industry".

Harry Cantrell, the General Electric representative on X3, reported that "IBM's activity in this standards work, its position papers, and its votes on various questions all point to a clear IBM intention to announce ASA-standard code equipment at an early date ... we can expect IBM's forthcoming 8000 line to be ASA compatible, featuring 8-bit bytes internally".

This heartening face of cooperation was welcomed by everyone, for ASCII was consonant in every aspect with the First Draft ISO Recommendation of 63 January [77](column 8). The ISO Draft did, however, maintain the 6-bit set with equal status, which was reflected in the adoption of ECMA-1, a 6-bit set [78] in March.

TC97/SC2 held its next meeting on 63 October, at which time it was decided to add the lower case alphabet. The Second Draft Recommendation [80] (column 10) gave more consideration to requirements for international alphabets; thus the single quote of the US gave way to accent acute, and the arrows fell to accent grave and circumflex, while underline was added. Format Effector 0 was definitely made Backspace, completing the necessary preparations for diacritical marks. One of the resolutions at this meeting contained the unanimous recommendation that this code be adapted for Numerical Control usage.

In preparation for the 3rd meeting of SC2, X3.2 was advised that it was wished to have the "at" symbol as an additional national use position, which implied a trade with underline. This was tentatively agreed. Also proposed was the replacement of RU by ACK, and ERR with NAK.

SC2 meeting No. 3 was held in New York City on 1964 May 12-15. It was relatively uneventful, except for the presence of a Japanese delegation of five, headed by Professor H. Wada. All of the proposed changes were approved, and ESC was moved in position. The revised proposal was to be circulated as the 3rd Draft [81](column 11) and sent to CCITT. Ad hoc groups were formed for implementation in media, but SC2 would not accept the bit-to-track assignment proposed by the US for magnetic tape, because they had not done any work on it (an 8-bit world was just beginning to seem real, with the announcement of the IBM 360).

The liaison with CCITT was significant. Its Study Group Special A had been considering a new alphabet since its 1958 meeting in Warsaw. At the 63 May meeting in Geneva, CCITT endorsed the principle of the 7-bit code for any new telegraph alphabet, and expressed general but preliminary agreement with the ISO work. It further requested the placement of the lower case alphabet in the unassigned area.

To match the 3rd Draft Proposal, X3.2 adopted Document X3.2.4/64 [83](column 12) on October 20, not knowing that CCITT had met in Moscow on 64 October 6-9 to study the ISO proposal as a working alphabet. CCITT made certain other changes for natural languages, replacing reverse slash by tilde and adding a universal currency symbol (column 13). At the request of the USSR (which did not participate in TC97/SC2), accent grave was interchanged with underline to accommodate the 31-character Cyrillic alphabet [96]. When advised of this, X3.2 accommodated with Document X3.2/172 [84] (column 14). In these last two revisions, however, added symbols for programming languages may be seen. In revising ASCII to remove all controls from column 7, Task Group X3.2.4 added (for the first time) the two braces, a vertical line, and a Logical Not symbol. The latter two were intended to accommodate the Logical Or and Not functions of PL/I. During the discussion, John Windhorst of CDC urged strongly that these two symbols be put somewhere in columns 2 through 5, in order to keep the PL/I 60-character subset within the dense graphic subset of 64 characters. His political wisdom did not prevail, and this led later to the unfortunate usage of exclamation point and circumflex for these functions, thus complicating an otherwise clean mapping to EBCDIC.

X3.2 being ever accommodating, X3.2/172 was superseded by Document X3.2/206 [86] (column 15) to match the 3-character symbol limitation of the new military standard [85]. This document carried an excellent summary of the various changes, with their justifications. However, it also carried a note that caused Univac to threaten a negative vote - "The resultant structure of specials (S), digits (D), and alphabetic (A) does not conform to the most prevalent collating convention (S-A-D) because of other more demanding code requirements". This despite my solution for passive logic in 1962, and despite the fact that it is impossible to provide satisfactory orderings of upper and lower case characters by simple binary comparison [47].

THE DARK AGES AGAIN

The position of IBM was a most important factor for progress of a standard code, and the System 360 was crucial to IBM's position. It was designed to handle both the Extended BCD Code [105] (for upward compatibility of much former equipment) and the eventual ASCII. However, the resistance in X3 and in ECMA to an 8-bit code, together with the fact that the ASCII printer and card reader were not ready when 360 announcement time neared, led to the decision to make EBCDIC the primary code. It was reasoned that ASCII could wait until the matter was settled, at which time the software would be modified slightly, the P-bit switched to ASCII internal mode, and everything would be fine.

Unfortunately, the software for the 360 was constructed by thousands of programmers, with great and unexpected difficulties, and with considerable lack of controls. As a result, the nearly $300 million worth of software (at first delivery) was filled with coding that depended upon the EBCDIC representation to work, and would not work with any other! Dr. Frederick Brooks, one of the chief designers of the IBM 360, informed me that IBM indeed made an estimate of how much it would cost to provide a reworked set of software to run under ASCII. The figure was $5 million, actually negligible compared to the base cost. However, IBM made the decision not to take that action, and from this time the worldwide position of IBM hardened to "any code as long as it is ours".

On 64 October 16, C.E.Mackenzie explained to GUIDE the plans of IBM for implementing ASCII in System 360. He stressed the profound difference between supporting the development of a standard and supporting the standard itself, such as actually implementing it in hardware and software. "IBM feels that its primary function is to supply equipment to meet the needs of the marketplace ... IBM will always make a businesslike decision ... the questionnaire distributed to all members of GUIDE at this meeting (and previously to SHARE) is one of the tools used to determine customer needs in the area of information interchange". He noted that the changes and differences between BCD and EBCDIC "should not provide any serious problems for IBM customers", and that "IBM has been unable to determine any appreciable customer needs for ASCII on magnetic tape, or on punched cards, or on perforated tape as input/output for a computer". He related the low need for data communication with ASCII.

In its 360 ESSG Information Letter No. 17 (64 October 14), IBM said that the choice between the two codes "is determined by a mode bit ... a sharp difference between the two codes is the collating sequence. The EBCDIC sequence is consistent with that of previous systems and is therefore largely (sic) compatible with that of our customers' files. The natural ASCII sequence, on the other hand, would place the numbers (sic) before the letters (not yet definitive)".

On this basis it was perhaps strange that it was J. Haddad of IBM who moved, at the 65 February 4 meeting, that X3 support the revised ASCII [86] for concurrent letter ballot and publication. Meanwhile, the Fourth ISO Draft Proposal [87](column 16), as influenced by CCITT, was being circulated as of March.

IBM gave X3 the results of its user surveys on April 22. Nevertheless the June 10 vote for prASCII (*p*roposed revised) stood at 31-0-0. IBM's affirmative ballot on the ISO proposal suggested the vertical line as logical OR, and the tilde as logical NOT. Auwaerter said that this had not been discussed in X3.2, as the comments arrived after all other work was done, and commitments made to the other countries. The US vote on the ISO Draft Proposal was affirmative, unanimously. At the 65 September 16 meeting, X3 voted 24-0-2 to forward the prASCII to ASA, encouraging the US Dept. of Commerce and the General Services Administration to solicit views on making ASCII a Federal standard under the provisions of the Brooks Bill, Public Law 89-306.

The voices of programmers grew louder, inversely with their understanding. Pressures for the two odd characters of the PL/I programming language continued, worldwide, and other pressures existed to put the "at" symbol back in the 64-character subset. Despite the express intent of many years, and rigorous intepretation of the meaning of "interchange code", one B. R. Faden proposed to SHARE, the IBM user group, that it should take a position with ASA for a study of an internal binary code. He said that "It would be a national calamity, truly national and truly calamitous, if ASCII is established as a standard for this purpose". This did not, however, prevent minicomputer designers from using ASCII as an internal code, and it was also chosen for the NCR Century series, the CDC 1800, and the Bull-GE 140.

A DRAFT RECOMMENDATION

TC97/SC2 held its 3rd meeting in Paris on 66 April 25-29, a momentous one. Japan presented Document 163 on extension to non-Latin alphabets, both 7-bit and an 8-bit superset [see 103]; USSR usage was reported as SO to Cyrillic, and SI to Latin [96]. To ensure full agreement, a joint meeting with CCITT, Working Group - Alphabet, was held concurrently. Commercial "at" and accent grave were once more interchanged, and the currency sign problem was alleviated, although not solved permanently, by assigning the pound sterling symbol to 2/3, with the number sign permissible for those countries that did not need it. SS (special sequence) was replaced by SUB (substitute character), and the overline (which could serve as a tilde) was put in 7/14. This produced ISO Draft Recommendation 1052 [90](column 17), for voting under accelerated procedures.

The US position was not stabilized by these actions. Knowing in advance of the proposal to interchange "at" and accent grave, the Information processing Systems Standards Board of ASA ordered suspension of printing of X3.2/206 as Standard X3.4-1965. The doors were in fact reopened, and at the next X3 meeting SHARE presented its views that the rASCII did not meet the pressing needs of the programming community, which was agreed by X3.4 also. However, X3.2 reported that it had listened to these arguments and still approved X3.2/384 (column 18), reflecting the changes of the Paris meeting. Following some editorial changes to allow programming people to stylize certain graphics for the logical operations, this document was sent for 30-day letter ballot.

Argument persisted, and at its 66 December 14 meeting X3.2 approved, in desperation, a modification to stylize the vertical bar by splitting it in two, making X3.2/411 (column 19). One of its members, Doug Kerr of Bell Telephone Labs, discovered in his research that, according to the encyclopaedia, the ASCII "are those inhabitants of the globe, which at certain times of the year, have no shadow ... such are the inhabitants of the torrid zone; by reason the sun is sometimes vertical to them". He observed that inhabitants of the parallel 41° 48′ N (running through such cities as ... Poughkeepsie, NY) are never ASCII. Yet when balloting closed, both IBM and API (American Petroleum Institute), carrying the campaign for SHARE, had voted affirmative. At the next X3 meeting it was forwarded to ASA by 33-0-2.

At TC97/SC2's 4th meeting, again in Paris, the vote on DR 1052 stood at 10-3 (2 not received) for the P (Principal) members of TC97. Spain voted negative because the tilde was in a national usage position, without full rank! The vote of the full ISO membership was 13-3-6 (32 not received). So it was sent to the ISO Council, with only Japan voting negative because of the dual assignment of pound sterling and number sign (the currency symbol remains a thorn to this day). The equivalent wording about stylizing for programmers was included. Thus it was a quiet meeting. For 15 resolutions times 10 countries there were 147 ayes, 1 nay, and 2 abstentions.

AN INTERNATIONAL STANDARD

On 67 December 22 the ISO accepted DR 1052 by a 12-1 vote (Ireland opposing), and it has been published as ISO Recommendation 646 [95]. It is entirely consistent with CCITT Working Alphabet No. 5. For the US it matches X3.4-1967 and -1968 [97] (put in force as a Federal Government standard by President Johnson). For Japan it is consistent with JISCII [103]. Many other countries have their national standards derived from it.

Primary emphasis today is not on changing the code. Work is in progress to propose *part* of an additional 128 positions to bring it to 256. This work originates primarily from the fact that the computer world is now predominantly using 8-bit characters on 9-track magnetic tape and other media. This work has been designated "expansion".

Conversely, even more effort has gone to "extension", which is the regularizing of reaching alternate 7-bit and even 8-bit sets via ESCape sequences and their registration. This is to make a harmonized structure wherein a multitude of important non-Latin symbol sets can be represented - Kata Kana, Kanji, Cyrillic, Arabic, meteorological symbols, bibliographic symbols, typesetting symbols and fonts - to mention only a few. It is this work that will demonstrate the vast significance of the establishment of an international standard code.

*Note: In the References that follow, the boldface **CR** followed by a number indicates that the number identifies the review of the referenced article in Computing Reviews, a publication of the Association for Computing Machinery.*

REFERENCES

Code (Articles and Papers)

1. A.Vandenburgh, "The Lincoln keyboard - a typewriter keyboard designed for computer input flexibility", *Commun. ACM* **1**, No. 7, 4 (1958 Jul).
2. R.F.Johnston, "Character representation and storage systems", *Proc. Canadian Conf. for Computing and Data Processing*, 120-135 (1958).
3. M.D.Grems, "A symbol coder for automatic documenting", *Comput. News*, No. 147, 9-18 (1959 Apr 15) and No. 148, 15-19 (1959 May 01).
4. H.S.Bright, "A proposed 64-character alphabet for hardware implementation of the IAL reference language", *Commun. ACM* **2**, No. 5, 6-9 (1959 May).
5. US Army Signal Corps, "Fieldata equipment intercommunication characteristics", 1959 Aug (revision of 1959 April).
6. J.Green, et al, "Remarks on ALGOL and symbol manipulation", *Commun. ACM* **2**, No. 9, 25-27 (1959 Sep).
   CR 00028
7. S.M.Newman, R.W.Swanson, K.C.Knowlton, "A notation system for transliterating technical and scientific texts for use in data processing systems", US Pat. Off. R&D Report No. 15, 1959.
8. R.W.Bemer, "A proposal for character code compatibility", *Commun. ACM* **3**, No. 2, 71-72 (1960 Feb).
   CR 00320
   [The ESCape character]
9. W.F.Luebbert, "Information handling and processing in large communication systems", Tech. Rep. 099-1, Stanford Electronic Labs, Stanford U. (1960 Jul 11).
10. E.A.Voorhees, "Some thoughts on reconciling various character set proposals", *Commun. ACM* **3**, No. 7, 408-409 (1960 Jul).
    [the 133-character set of the Los Alamos STRETCH computer]
11. H.J.Smith, Jr., "A short study of notation efficiency", *Commun. ACM* **3**, No. 8, 468-473 (1960 Aug).
    [presages the 8-bit byte of the IBM 360]
12. R.W.Bemer, F.A.Williams, "A survey of coded character representation", *Commun. ACM* **3**, No. 12, 639-641 (1960 Dec).
    CR 00639
13. H.McG.Ross, "Considerations in choosing a character code for computers and punched tapes", *The Computer J.* **3**, No. 4, 202-210 (1961 Jan).
14. R.W.Bemer, H.J.Smith, Jr., F.A.Williams, "Design of an improved transmission/data processing code", *Commun. ACM* **4**. No. 5, 212-217, 225 (1961 May).
15. A.A.Markov, "On alphabet coding", *Doklady Akad. Nauk* **139**, No. 3, 560-561 (1961 Jul).
    CR 03662 (null)
16. R.R.Shaw, "Parameters for machine handling of alphabetic information", *Amer. Soc.* **13**, No. 3, 267-269 (1962 Jul).
    CR 03925
    [claims 6-bit codes are insufficient, need 11-bit. No exposure to extension procedures]
17. R.F.Clippinger, "Data processing standards", Computer Applications Symposium, Armour Res. Found., 1962, 176-181.
18. C.A.Phillips, R.E.Utman, "Standardization in computers and information processing", *Proc. Fall Joint Comput. Conf.*, 1962, 177-183
19. R.W.Bemer, W.W.Buchholz, "Character set", Chapter 6, *Planning a Computer System*, McGraw-Hill, 1962, 60-74.
20. H.McG.Ross, "Current positions on standards work relating to computers", *The Comput. Bull.* **6**, No. 4, 133-135 (1963 Mar).
    CR 04658
    [says First ISO draft primarily a synthesis of BSI and ASA proposals]
21. T.Merrill, "Versatile passkey to communication", *Business Week Magazine*, 1963 Jun 15, 128-131.
22. -, "New ASA standard offers coding uniformity for data interchange", Special Issue, *IBM Data Processor* **7**, No. 7, 12 pp. (1963 Aug 07).
23. "Code standardization", Report for Aerospacecom, Deputy for Communication Systems, by ITT Commun. and Systems, 1963 Aug 15.
24. R.W.Bemer, "The American standard code for information interchange", Parts 1 and 2, *Datamation* **9**, No. 8, 32-36 (1963 Aug) and **9**, No. 9, 39-44 (1963 Sep).
    CR 04910
25. L.L.Griffin, et al, "Uniform machine language to speed communications", *The Magazine of Standards* (ASA) **34**, No. 8, 235-239 (1963 Aug).
26. E.H.Clamons, "At last - a standard code for information systems", *Control Engineering* **10**, No. 10, 91-94 (1963 Oct).
27. -, "American standard code for information interchange", *Commun. ACM* **6**, No. 8, 422-426 (1963 Aug).
    CR 04943
    ["no collating sequence is specified"]
28. -, "ALCOR group representation of ALGOL symbols", *Commun. ACM* **6**, No. 10, 597-599 (1963 Oct).
    CR 05697
29. J.F.Auwaerter, "A new standard code for teletypewriters", *Bell Laboratories Record*, 1963 Nov, 395-400.
30. -, "Preliminary standards: data processing, representation of ALGOL symbols", *Math. Tech. Wirts.* **3**, No. 10, 117-119 (1963).
    CR 05362
    [German DIN standard for cards and 5-track perforated tape]
31. -, "Report on CCITT data communications study group meeting", *Commun. ACM* **7**, No. 3, 152 (1964 Mar).
32. C.J.Shaw, "On declaring arbitrarily coded alphabets", *Commun. ACM* **7**, No. 5, 288-290 (1964 May).
    CR 06672
    [notes the inability of FORTRAN and ALGOL to handle alphabetic data, while COBOL is limited to a single machine-dependent alphabet. The advent of ASCII should give impetus to ability to declare a specific alphabet and/or collating sequence]
33. E.Bukstein, "Binary computer codes and ASCII", *Electronics World*, 1964 Jul, 28-29.
34. V.E.Henriques, "Current position and future of EDP standards", *Proc. ACM 19th Natl. Conf.*, 2 pp, 1964 Aug.
    CR 07095
35. V.E.Henriques, "The urgent quest for standards", *Business Autom.* **11**, No. 9, 22-27 (1964 Sep).
    CR 08438
36. -, "The BCS standards advisory committee", *The Comput. Bull.* 8, No. 2, 61-63 (1964 Sep).
37. E.Morenoff, J.B.McLean, L.Odell, "IPC, a coded character set for information processing", Rome Air Dev. Center, Griffiss AFB, AD-608916, 1964 Oct 18, 18 pp.
    [8-bit set, subsettable to 7, 6, 5, 4].
38. H.McG.Ross, "The I.S.O. character code", *The Comput. J.* **7**, No. 3, 197-202 (1964 Oct).
    CR 07488
39. D.W.Barron, "The ISO character code", *The Comput. J.* **7**, No. 4, 281 (1965 Jan).
    CR 11206
    [A programmer comments that "standardization of character codes is not as important as sometimes made out. Computers can translate, can't they?]
40. -, "ACM opposes acceptance of EIA standard RS-244 as an American standard on a par with ASCII", *Commun. ACM* 8, No. 2, 102 (1965 Feb).
    CR 08543
41. -, "Proposed revised American standard code for information interchange", *Commun. ACM* 8, No. 4, 207-214 (1965 Apr).
    CR 07784
    [X3.2/206, 1965 Jan 21 - see ref. 86].
42. H.Zemanek, "Alphabete und Codes", *Electronische Rechenanlagen* 7, No. 5, 239-258 (1965 Oct)
43. W.S.Crosby, "Note on an ASCII-octal code table", *Commun. ACM* **8**, No. 10, 616 (1965 Oct).
    CR 08765 (null)
44. L.Davidson, "A pushbutton telephone for alphameric input", *Datamation* **12**, No. 4, 27-30 (1966 Apr).
    CR 12922

45. R.W.Bemer, "Information processing - a bit expensive?", Forum Editorial, *Datamation* 12, No. 6, 140 (1966 Jun).
46. -, "Code extension in ASCII (an ASA tutorial) ", *Commun. ACM* 9, No. 10, 758-762 (1966 Oct).
47. R.W.Bemer, "ESC facility in ASCII", letter, *Commun. ACM* 10, No. 4, 282 (1967 Apr).
48. M.E.Stevens, J.L.Little, "Automatic typographic-quality typesetting techniques: a state-of-the-art review", *NBS Monograph 99*, 98 pp, 1967 Apr.
CR 13088
["discusses problems of character sets that are adequate, computer programming, keyboard design"]
49. H.Zemanek, "Gedanken zu einem 8-bit-code", *Electronische Rechenanlagen* 9, 1967 Apr, 65-67.
50. H.Berndt, G.Haberzettl, "Der internationale 7-bit-code", *Electronische Rechenanlagen* 9, 1967 Apr, 68-73.
51. M.H.Lewin, "Portable electronic keyboard for computer input by telephone", *IEEE Trans.* ECU3-16, No. 3, 332-334 (1967 Jun).
CR 12924
["generates all 128 codes"]
52. D.L.A.Barber, D.O.Claydon, "A keyboard layout for use with the ISO seven bit code", *Comput. Bull.* 11, No. 1, 66-67 (1967 Jun).
53. P.B.Goodstat, "USASCII, What's it all about", *Data Processing Magazine*, 1967 Jun, 20-24.
54. A.M.Bondarev, "Input and output devices for electronic computers", AD-662762, 67 Jul 28, 8 pp.
[translation of *Standartizatsiya* 3, 1964; references the USSR standard GOST-10525-63].
55. R.W.Bemer, "Towards standards for handwritten zero and oh", *Commun. ACM* 10, No. 8, 513-518 (1967 Aug).
CR 14629
56. -, "Proposed USA standard - magnetic tape labels for information interchange", *Commun. ACM* 10, No. 11, 737-743 (1967 Nov).
CR 13712
57. -, "Proposed USA standard - general purpose alphanumeric keyboard arrangement for information interchange", *Commun. ACM* 11, No. 2, 126-129 (1968 Feb).
58. D.V.Black, "Creation of computer input in an expanded character set", *J. Library Autom.* 1, No. 2, 110-120 (1968 Jun).
CR 15755
["101 graphics - less detailed and complex than the recently established MARC II standard"].
59. J.N.Bairstow, "USASCII: Federal Government action on standards challenges computer manufacturers", *Control Engineering*, 1968 Aug, 84-87.
60. P.Hirsch, "Release due for Federal directive on ASCII", News Scene, *Datamation* 14, No. 10, 69 (1968 Oct).
61. -, "Correspondences of 8-bit and Hollerith codes for computer environments - a USASI tutorial", *Commun. ACM* 11, No. 11, 783-789 (1968 Nov).
CR 16119
[reconciles 8-bit ASCII with 8-bit EBCDIC by mapping each to the other via Hollerith assignments]
62. -, "Proposed USA standard - code extension procedures for information interchange", *Commun. ACM* 11, No. 12, 849-852 (1968 Dec).
63. -, "Standards in information processing", *Elektronische Datenverarbeitung* 10, No. 9, 458-459 (1968).
CR 17173
64. G.E.Jones, "The impact of standards", *Computers and Automation* 18, No. 5, 38-39 (1969 May).
CR 17385
["development expense is substantial ... standardization should be preceded by clear statement of need, and de facto practices should be incorporated wherever possible"].
65. J.L.Little, "Some evolving conventions and standards for character information coded in six, seven, and eight bits", Natl. Bur. Stds., NBS-TN-478, 1969 May, 33 pp.
66. C.E.Cohn, "Reading ASCII tape on a BCD computer", *Software Age*, 1969 Jul.
67. R.W.Bemer, "Escape to reality", Forum Editorial, *Datamation* 15, No. 8, 239-240 (1969 Aug).
68. J.P.Ancona, S.M.Garland, J.J.Tropsa, "At last: a standard for keyboards", *Datamation* 17, No. 5, 32-36 (1971 Mar 01).

Code (Drafts and Standards)

69. -, "Fieldata equipment intercommunication characteristics", Memorandum for Director, Data Proc. Facil. Div., Commun. Dept., USASRDL, 1959 Apr (rev. 1959 Aug 01).
70. R.W.Bemer, W.W.Buchholz, "An extended character set standard", IBM Tech. Publ. TR 00.721, 1960 Jun (revision of TR 00.18000.705, 1960 Jan).
CR 00813
[as in Stretch].
71. Electronic Industries Assoc., "Basic character set code, tentative standards proposal", TR24.4 Doc. 7233 (1960 May).
[language and format for punched and magnetic tapes, incorporating ESC].
72. US Dept. of Defense, "Fieldata code", in MIL-STD-188A, 1960 Jun 21.
73. X3.2/1, "Proposed American standard code for information interchange", 1961 Nov 10.
74. ISO/TC97/WG B(Sec.-5)16, 1962 Jun.
[proposals A and B].
75. X3.2/4, Proposed American standard code for information interchange", 1962 May 25.
[to meet AIEE requirements].
76. IBM Document J22-6690, "Standard BCD interchange code", 1962.
["defines for the IBM Corporation a standard 64-character set for the IBM 1401, 1410, 7040, and 7044 Data Processing Systems". Not for the 709, 705-7080 series, or any other equipment].
77. ISO/TC97/SC2(Sec.-12)30, "ISO draft proposal, 6 and 7 bit character codes for information processing interchange", 1963 Jan.
78. ECMA-1, "ECMA standard for a 6 bit input/output character code", 1963 Mar 29.
79. ANS X3.4-1963, "American standard code for information interchange", 1963 Jan 17.
[was X3.2/13, 1962 Sep 10].
80. ISO/TC97/SC2(Sec.-22)75, "Second ISO draft proposal, 6 and 7 bit character codes for information processing interchange", 1963 Dec.
81. ISO/TC97/SC2(Sec.-32)125, "Third ISO draft proposal, 6 and 7-bit character codes for information processing interchange", 1964 May.
82. 7-bit code, CCITT version, 1964 Oct 09.
83. X3.2.4/64, "Draft revised ASCII", 1964 Oct 20.
84. X3.2/172, "Draft proposed revised ASCII", 1964 Dec 03.
85. M-64-900A, "Mil Std Code", 1964 Dec 01.
86. X3.2/206, "Proposed revised ASCII", 1965 Jan 21.
87. ISO/TC97/SC2(Sec.-37)130, "Fourth ISO draft proposal, 6 and 7 bit character codes for information processing interchange", 1965 Mar.
88. ECMA-6, "ECMA standard for a 7 bit input/output character code", 1965 Apr.
[1st edition, superseding drafts TC1/67/5 and TC1/66/69]
89. X3.2/384, "Draft proposed revised ASCII", 1966 May 20.
90. ISO Draft Recommendation 1052, "6 and 7 bit character codes for information processing interchange", 1966 Jun.
91. DIN 66 003, "Informationsverarbeitung, 7-bit-code", 1966 May.
92. X3.2/411, "Draft proposed revised ASCII", 1966 Jul 15.
93. ECMA-6, "ECMA standard for a 7 bit input/output character code", 1967 Jun (2nd edition).
94. ECMA-14, "ECMA standard for rules for the definition of 4 bit sets derived from the ECMA 7 bit coded character set for information interchange", 1967 Nov.
95. ISO Recommendation 646, "6 and 7 bit coded character sets for information processing interchange", 1967 Dec (DR 1052).
96. USSR State Standard, GOST 13052-67, "Computers and data transmission equipment, alpha-numeric codes", approved 1967 Jul 10. [mandatory, by law].
97. X3.4-1968, "USA standard code for information interchange". [supersedes X3.4-1967 (equivalent to FIPS 1, signed by President Lyndon Johnson on 1968 May 11) by adding NL. Approved 1968 Oct 10].
98. ECMA-17, "ECMA standard for graphic representation of control characters of the ECMA 7 bit coded character set for information interchange", 1968 Nov.
99. ISO Recommendation 963, "Guide for the definition of 4-bit character sets derived from ISO 7-bit coded character set for information processing", 1969 Feb (DR 1322).

100. ECMA-19, "ECMA standard for the coding of character sets for MICR and OCR", 1969 Jun.
101. "Proposed revision of ECMA-6", (3rd edition, 1969, TC1/69/13).
102. ECMA-23, "ECMA standard for keyboards generating the code combinations of the characters of the ECMA 7 bit coded character set", 1969 Jun.
103. Japanese Industrial Standard Code for Information Interchange, "JIS-CII", C6220-1969.
104. ECMA-6, "ECMA standard for a 7 bit input/output character code", 1970 Jul (3rd edition).
105. IBM Corporate Systems Standard, "Extended BCD interchange code, Latin alphabets", CSS 3-3220-002, 1970 Nov, 40 pp.
106. ECMA-35, "ECMA standard for extension of the 7-bit coded character set", 1971 Dec.
107. ISO Draft Recommendation 1672, "Hardware representation of ALGOL basic symbols in the ISO 6 and 7 bit coded character sets".
108. ISO Draft Recommendation 2022, "Code extension procedures for ISO 7 bit code".
109. ISO Draft Recommendation 2033, "Coding of character sets for MICR and OCR".
110. ISO Draft Recommendation 2047, "Graphical representation for the control characters of the ISO 7-bit character set".
111. ISO Draft International Standard, DIS 2375, "Data processing - Procedures for the registration of escape sequences". (1970 Sep).

Magnetic Tape (Articles and Papers)

112. R.R.Bender, D.J.Galage, "Packing mode control", IBM Tech. Disclosure Bull. 4, 1961 Aug, 61-63.
[packed numerics of an 8-bit code on magnetic tape].
113. -, "Proposed American standard - recorded magnetic tape for information interchange", *Commun. ACM* 7, No. 5, 284-286 (1964 May).
[X3.2/73, 64 Jan 31]
CR 06471
114. R.W.Bemer, "Information processing - a bit expensive?", Forum Editorial, *Datamation* **12**, No. 6, 140 (1966 Jun).
115. -, "Proposed USA standard - recorded magnetic tape for information interchange (200 cpi, nrzi)", *Commun. ACM* **10**, No. 11, 730-737 (1967 Nov).
CR 13711
116. -, "Proposed USA standard - magnetic tape labels for information interchange", *Commun. ACM* **10**, No. 11, 737-743 (1967 Nov).
CR 13712

Magnetic Tape (Drafts and Standards)

117. ANS X3.22-1967, "Recorded magnetic tape for information interchange (800 cpi, NRZI) ", (equivalent to FIPS 3).
118. ECMA-12, "ECMA standard for data interchange on 9 track magnetic tape at 31,5 bit per mm (800 bpi) ", 1967 Nov (2nd edition 1970 Jun).
119. ISO Recommendation 961, "Implementation of the 6 and 7 bit coded character sets on 7 track 12.7 mm (1/2 in.) magnetic tape", 1969 Feb (DR 1320).
120. ISO Recommendation 962, "Implementation of the 7 bit coded character set on 9 track 12.7 mm (1/2 in.) magnetic tape", 1969 Feb (DR 1321).
121. ECMA-5, "Data interchange on 7-track magnetic tape", 1970 Jun (3rd edition).
122. ISO Recommendation 1861, "7-track 8 rpmm (200 rpi) magnetic tape for information interchange", 1971 May (DR 1861).
123. ISO Recommendation 1862, "9-track 8 rpmm (200 rpi) magnetic tape for information interchange", 1971 May (DR 1872).
124. ECMA-34, "ECMA standard for data interchange on 3,81 mm magnetic tape cassette (32 bpmm, phase encoded) ", 1971 Sep.
125. ECMA-36, "ECMA standard for data interchange on 9-track phase encoded magnetic tape at 63 bpmm (1600 bpi) ", 1971 Dec.

Punch Cards (Articles and Papers)

126. R.W.Bemer, "A proposal for a generalized card code for 256 characters", *Commun. ACM* **2**, No. 9, 19-23 (1959 Sep).
CR 00025
127. British Standards Institution, "Alphanumeric punching codes for data processing cards", Standard BS3174:1959.
128. R.W.Bemer, W.W.Buchholz, "An extended character set standard", IBM Tech. Publ. TR 00.721, 1960 Jun (revision of TR 00.18000.705, 1960 Jan).
129. H.J.Smith, Jr., F.A.Williams, "Survey of punched card codes", *Commun. ACM* **3**, 638, 642 (1960 Dec).
130. D.M.Grant, "A new trinary code compared with existing codes for punched cards", *J. Chem. Doc.* **1**, No. 1, 68-69 (1961 Jan).
CR 02713 (null)
131. H.McG.Ross, "Further survey of punched card codes", *Commun. ACM* 4, No. 4, 182-183 (1961 Apr).
132. -, "ECMA standard - representation of the standard ECMA 7-bit code in punched cards", *Commun. ACM* **8**, No. 7, 441-442 (1965 Jul).
[so-called Decimal ASCII]
CR 09937
133. E.H.Clamons, "The proposed new punched card code standard", *Control Engg.* **14**, No. 2 (1967 Feb).
134. -, "Proposed USA standard - Hollerith punched card code", *Commun. ACM* 11, No. 4, 275-281 (1968 Apr).
135. -, "Correspondences of 8-bit and Hollerith codes for computer environments - a USASI tutorial", *Commun. ACM* **11**, No. 11, 783-789 (1968 Nov).
CR 16119

Punch Cards (Draft and Standards)

136. British Standards Institution, "Draft British standard for codes for punched cards", CZ(MEE)6770, 1959 May.
137. EIA RS-292, "Punched Card Standard".
138. ANS X3.26-1969, "Hollerith punched card code" (same as FIPS 14).
[128 characters only]
139. ECMA-20, "ECMA standard for the implementation of the ECMA 7 bit coded character set on punched cards", 1969 Jun.
140. ECMA-25, "ECMA standard for representation of 8 bit combinations in 12 row punched cards", 1970 Aug.
[superseded draft TC1/69/78, 1969 Dec]
141. ANS X3.26-1970, "Hollerith punched card code".
[256 characters - revision of X3.26-1969]
142. ISO Recommendation 1679, "Representation of ISO 7 bit coded character set on 12 row punched cards", 1970 Jul (DR 1679).
143. ISO Recommendation 2021, "Representation of ISO 8 bit coded character set on 12 row punched cards", 1971 May (DR 2021).

Perforated Tape (Articles and Papers)

144. -, "Proposed American standard - perforated tape code for information interchange", *Commun. ACM* 7, No. 6, 336-338 (1964 Jun).
[X3.2/16, 63 Dec 19]
CR 06517

Perforated Tape (Drafts and Standards)

145. Electronic Industries Assoc., "Basic character set code, tentative standards proposal", TR24.4 Doc. 7233 (1960 May).
146. British Standards Institution, AA(DPE)3543, "Draft British standard for punched tape coding, part 1, 7 track code", (1960 Sep), (eventually BS3480:1962).
147. ANS X3.6-1965, "Perforated tape code for information interchange", (approved 1965 Jul 09, equivalent to FIPS 2).
148. ISO Recommendation 1113, "Representation of 6 and 7 bit coded character sets in punched tape", (1969 Nov), (DR 1418).
149. ECMA-10, "ECMA standard for data interchange on punched tape", 1970 Jul (2nd edition).

Communications (Articles and Papers)

150. -, "Proposed American standard - bit sequencing of the American standard code for information interchange (ASCII) in serial-by-bit data transmission", *Commun. ACM* 7, No. 6, 333-336 (1964 Jun).
[X3.3/16]
CR 06516
[expository remarks say that conclusion for HOF is based upon economical design of equipment; contains complete summary of input documents].

151. F.C.White, "Bit sequence stirs data/comm feud", *Air Transport World*, 1964 Jul, 26-28.

152. R.W.Bemer, J.Booth, "Comments on bit-sequencing of the ASCII in serial-by-bit data transmission", *Commun. ACM* 7, No. 8, 483-485 (1964 Aug).
CR 06519
["offers powerful arguments in favor of 'low-order' bit sequencing, which is contrary to the proposed American standard"].

153. -, "Transparent-mode control procedures for data communication, using the American standard code for information interchange - a tutorial", *Commun. ACM* **8**, No. 4, 203-206 (1965 Apr).
CR 07783

154. E.H.Clamons, "Low-order or high-order, ?tsrif tib hcihw", *Control Engineering*, 1965 Aug, 102-103

155. -, "Proposed American standard, character structure and character parity sense for serial-by-bit data communication in the ASCII", *Commun. ACM* 8, No. 9, 553-556 (1965 Sep).
[X3.3/23].

156. -, "Proposed American standard, character structure and character parity sense for parallel-by-bit data communication in the ASCII", *Commun. ACM* **9**, No. 9, 695-697 (1966 Sep).
[X3.3/35[.

157. -, "Data communication control procedures for the USA standard code for information interchange", *Commun. ACM* **12**, No. 3, 166-178 (1969 Mar).
[same as ISO/TC97/SC6/278].
CR 16858

158. W.L.Harper, "The remote world of digital switching", *Datamation* **17**, No. 6, 22-31 (1971 Mar 15).

159. S.R.Rosenblum, "Progress in control procedure standardization", *Proc. 2nd Symp. on Problems in the Organization of Data Communication Systems*, 153-159 (1971).
CR 22425

Communications (Drafts and Standards)

160. X3.3.3/7, "Working paper, character structure for serial-by-bit transmission in the ASCII", 1964 Jul 07.
[first separation of parity sense from bit sequence].

161. ANS X3.15-1966, "Bit sequencing of the ASCII in serial-by-bit data transmission" (equivalent to FIPS 16).
[X3.3/23, 65 Jun 09, approved 66 Aug 19].

162. ANS X3.16-1966, "Character structure and character parity sense for serial-by-bit data communication in the ASCII" (equivalent to FIPS 17).
[X3.3/..., approved 66 Aug 19. ODD for synchronous, EVEN for asynchronous].

163. ECMA-16, "ECMA standard for basic mode control procedures for data communications systems using the ECMA 7 bit code", 1968 May.

164. ANS X3.25-1968, "Character structure and character parity sense for parallel-by-bit data communication in the ASCII" (equivalent to FIPS 18).
[X3.3/35, 66 Feb 09, approved 68 Oct 24].
[equivalent to ISO/TC97/SC6(USA-15)124].

165. ECMA-24, "ECMA standard for code independent information transfers", 1969 Dec.

166. ECMA-26, "ECMA standard for recovery procedures", 1971 Apr.

167. ECMA-27, "ECMA standard for abort and interrupt procedures", 1971 Apr.

168. ECMA-28, "ECMA standard for multiple station selection procedures", 1971 Apr.

169. ECMA-29, "ECMA standard for conversational information transfer", 1971 Sep.

170. ANS X3.28-1971, "Procedures for the use of the communication control characters of the ASCII in specified data communication links".

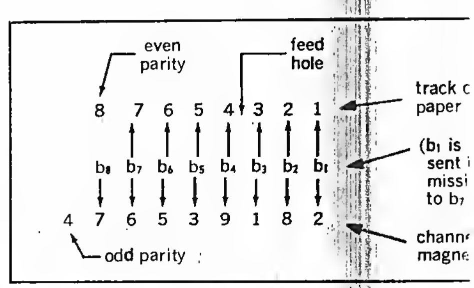
171. ISO Draft International Standard, DIS 2111, "Code independent transmission procedures".

3-char control (7-bit code)	alternate 2-char mnemonic	ANSI mnemonic (8-bit code)	ECMA mnemonic (if different)	
ACK	(AK)			Acknowledge
BEL	(BL)			Bell
BS				Backspace
CAN	(CN)			Cancel
		CD	DCL	Character Delete
		CI	ICL	Character Insert
		CIF		Character Insert Off
		CIN		Character Insert On
		CLC	EEL	Clear Line from Cursor
CR				Carriage Return
		CSC	EED	Clear Screen from Cursor
DC1	(D1)			Device Control 1
DC2	(D2)			Device Control 2
DC3	(D3)			Device Control 3
DC4	(D4)			Device Control 4
DEL	(DT)			Delete
DLE	(DL)			Data Link Escape
EM				End of Medium
ENQ	(EQ)			Enquiry
		EO		Eight Ones
EOT	(ET)			End of Transmission
ESC	(EC)			Escape
		ESI		Extended Shift In
		ESO		Extended Shift Out
ETB	(EB)			End of Transmission Block
ETX	(EX)			End of Text
FF				Form Feed
FS				File Separator
GS				Group Separator
		HF	RPM	Highlight Off
		HLF	FHL	Half Line Feed
		HLR	RHL	Half Line Reverse Feed
		HN	APM	Highlight On
HT				Horizontal Tab
		HTC		Horizontal Tab Clear
		HTS		Horizontal Tab Set
		LCF		Local Copy Off (full duplex)
		LCN		Local Copy On (half duplex)
		LD	DL	Line Delete
LF				Line Feed
		LI	IL	Line Insert
NAK	(NK)			Negative Acknowledge
		NP		Next Page
NUL	(NU)			Null
		PD	CUD	Cursor (Pointer) Down
		PFF		Protect Format Off
		PFN		Protect Format On
		PH	CUH	Cursor (Pointer) Home
		PL	CUB	Cursor (Pointer) Left
		PM		Cursor (Pointer) Return
		PP		Previous Page
		PR	CUF	Cursor (Pointer) Right
		PT		Cursor (Pointer) Tab
		PU	CUU	Cursor (Pointer) Up
		RLF		Reverse Line Feed
RS				Record Separator
		SD		Scroll Down
SI				Shift In
SO				Shift Out
SOH	(SH)			Start of Heading
SP				Space (a blank)
STX	(SX)			Start of Text
		SU		Scroll Up
SUB	(SB)			Substitute
SYN	(SY)			Synchronous Idle
US				Unit Separator
VT				Vertical Tab
		VTC		Vertical Tab Clear
		VTS		Vertical Tab Set

Note 1
These 12 positions are variable — 2 for currency, 7 primary national usage, and 3 secondary usage which are diacritical marks used for alphabetical extension when preceded by BS. Positions 2/7 and 2/12 are invariant but also serve as diacritical marks. The presently-known assignments are given in the table below.

	currency		1st 7 national				dia	dia	1st 7 national			dia
	2/3	2/4	4/0	5/11	5/12	5/13	5/14	6/0	7/11	7/12	7/13	7/14
Netherlands—A			@	[	\	]	^	`	{	|	}	~
Australia	#											
Belgium—A												
W. Germany—A		$										
US												
Japan					¥							
UK	£				\		↑					
Italy—A	#	¤					^					
Switzerland—A		¤					^					
France—A		$										
USSR		¤			V		^	`				
Netherlands—B					IJ					ij		
Belgium—B			à		ç				é	ij	è	¨
France—B	£	$	à	°	ç	§	^	`	é	ù	è	¨
Switzerland—B			à		ç				é	ù	è	¨
Italy—B	#	$	§	°	ç	é	^	ù	à	ò	è	ì
Switzerland—C					ç	é		ù	à	è	é	ì
Hungary	#	Ft	§	É	Ö	Ü	Á	ó	é	ö	ü	á
W. Germany—B	£	$	§	Ä	Ö	Ü	^	`	ä	ö	ü	β
Switzerland—D				Ä	Ö	Ü			ä	ö	ü	β
Sweden	#	¤	É	Ä	Ö	Å	Ü	é	ä	ö	å	ü
Finland				Ä	Ö	Å			ä	ö	å	
Denmark				Æ	Ø	Å			æ	ø	å	
Norway	¤			Æ	Ø	Å			æ	ø	å	
Spain					Ñ					ñ		

$b_8 b_7 b_6 b_5$			0000	0001	0010
	$b_4 b_3 b_2 b_1$	ROW \ COL	0	1	2
	0000	0	NUL 12-0-9-8-1	DLE 12-11-9-8-1	SP
	0001	1	SOH 12-9-1	DC1 11-9-1	! 12
	0010	2	STX 12-9-2	DC2 11-9-2	"
	0011	3	ETX 12-9-3	DC3 11-9-3	NOTE 1 £ / #
	0100	4	EOT 7-9	DC4 4-8-9	NOTE 1 $ / ¤ 11
	0101	5	ENQ 0-9-8-5	NAK 9-8-5	% 0
	0110	6	ACK 0-9-8-6	SYN 9-2	& 12
	0111	7	BEL 0-9-8-7	ETB 0-9-6	'
	1000	8	BS 11-9-6	CAN 11-9-8	(12
	1001	9	HT 12-9-5	EM 11-9-8-1	) 11
	1010	10	LF 0-9-5	SUB 9-8-7	* 11
	1011	11	VT 12-9-8-3	ESC 0-9-7	+ 12
	1100	12	FF 12-9-8-4	FS 11-9-8-4	, 0
	1101	13	CR 12-9-8-5	GS 11-9-8-5	- 11
	1110	14	SO 12-9-8-6	RS 11-9-8-6	. 12
	1111	15	SI 12-9-8-7	US 11-9-8-7	/ 0



—— 3-char control (7-bit code)
—— alternate 2-char mnemonic
—— ANSI mnemonic (8-bit code)
—— ECMA mnemonic (if different)

Acknowledge
Bell
Backspace
Cancel
DCL Character Delete
ICL Character Insert
Character Insert Off
Character Insert On
EEL Clear Line from Cursor
Carriage Return
EED Clear Screen from Cursor
Device Control 1
Device Control 2
Device Control 3
Device Control 4
Delete
Data Link Escape
End of Medium
Enquiry
Eight Ones
End of Transmission
Escape
Extended Shift In
Extended Shift Out
End of Transmission Block
End of Text
Form Feed
File Separator
Group Separator
RPM Highlight Off
FHL Half Line Feed
RHL Half Line Reverse Feed
APM Highlight On
Horizontal Tab
Horizontal Tab Clear
Horizontal Tab Set
Hard Copy Off (full duplex)
Hard Copy On (half duplex)
DL Line Delete
Line Feed
IL Line Insert
Negative Acknowledge
New Page
CUD Cursor (Pointer) Down
Format Off
Format On
CUH Cursor (Pointer) Home
CUB Cursor (Pointer) Left
Cursor (Pointer) Return
Previous Page
CUF Cursor (Pointer) Right
Cursor (Pointer) Tab
CUU Cursor (Pointer) Up
Reverse Line Feed
Record Separator
Down
Shift In
Shift Out
Start of Heading
Space (blank)
Start of Text
Up
Substitute
Synchronous Idle
Unit Separator
Vertical Tab
Vertical Tab Clear
Vertical Tab Set

JISCII (Japanese Industrial Standard Code for Information Interchange) is an 8-bit code consisting of the ISO characters plus the Kata Kana characters shown in the upper row positions of columns 10-13 (columns 8 and 9 are used for additional controls, 14 and 15 for additional graphics).

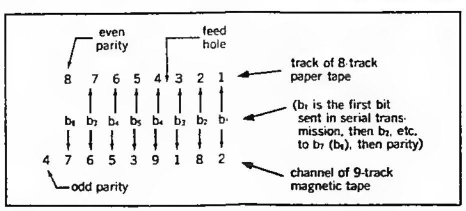
GOST 13052-67 defines the USSR set shown in the lower row entry positions of columns 12-15. Actually, the standard defines these characters for columns 4-7 (SO = Cyrillic register, SI = Latin register). Columns 8-11 are identical to 0-3.

(ESC) (CHAR.) — SOFT-COPY CONTROLS — (SINGLE CHAR.)

$b_8\,b_7\,b_6\,b_5$		0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$b_4\,b_3\,b_2\,b_1$	COL / ROW	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0000	0	NUL □	DLE	SP △	0	NOTE 1 @ / CI	P / CD	NOTE 1	p	CI	CD		ー	タ / ю	ミ / п	Ю	П
0001	1	SOH	DC1	!	1 / HTS	A / PU	Q / CIN	a	q	PU	CIN	。	ア	チ / а	ム / я	А	Я
0010	2	STX ⊥	DC2	"	2 / HTC	B / PD	R / CIF	b	r	PD	CIF	「	イ	ツ / б	メ / р	Б	Р
0011	3	ETX	DC3	NOTE 1 £/#	3 / HN	C / PR	S / SU	c	s	PR	SU	」	ウ	テ / ц	モ / с	Ц	С
0100	4	EOT	DC4	NOTE 1 $/¤	4 / HF	D / PL	T / SD	d	t	PL	SD	、	エ	ト / д	ヤ / т	Д	Т
0101	5	ENQ ⊠	NAK	%	5 / VTS	E	U / NP	e	u		NP	・	オ	ナ / е	ユ / у	Е	У
0110	6	ACK ✓	SYN	&	6 / VTC	F	V / PP	f	v		PP	ヲ	カ	ニ / ф	ヨ / ж	Ф	Ж
0111	7	BEL	ETB	'	7 / RLF	G / PM	W / PFN	g	w	PM	PFN	ァ	キ	ヌ / г	ラ / в	Г	В
1000	8	BS	CAN ⌛	(	8 / HLR	H / PH	X / PF	h	x	PH	PF	ィ	ク	ネ / х	リ / ь	Х	Ь
1001	9	HT →	EM	)	9 / HLF	I / PT	Y	i	y	PT		ゥ	ケ	ノ / и	ル / ы	И	Ы
1010	10	LF ≡	SUB	*	: / LCF	J / CSC	Z	j	z	CSC		ェ	コ	ハ / й	レ / з	Й	З
1011	11	VT ↓	ESC ⊖	+	; / LCN	K / CLC	NOTE 1 [	k	NOTE 1 {	CLC		ォ	サ	ヒ / к	ロ / ш	К	Ш
1100	12	FF	FS	,	<	L / LI	NOTE 1 \	l	NOTE 1 ¦	LI		ャ	シ	フ / л	ワ / э	Л	Э
1101	13	CR ←	GS	–	=	M / LD	NOTE 1]	m	NOTE 1 }	LD		ュ	ス	ヘ / м	ン / щ	М	Щ
1110	14	SO ⊗	RS	.	>	N	NOTE 1 ^	n	NOTE 1 ‾			ョ	セ	ホ / н	゛ / ч	Н	Ч
1111	15	SI ⊙	US	/	?	O	_	o	DEL			ッ	ソ	マ / о	゜ / _	О	EO

Note 1
These 12 positions are variable — 2 for currency, 7 primary national usage, and 3 secondary usage which are diacritical marks used for alphabetical extension when preceded by BS. Positions 2/7 and 2/12 are invariant but also serve as diacritical marks. The presently-known assignments are given in the table below.

currency		1st 7 national			dia	dia	1st 7 national			dia
2/4	4	5/11	5/12	5/13	5/14	6/0	7/11	7/12	7/13	7/14
$		[	\	]	^	`	{	¦	}	‾
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]



Alternate controls in these 5 columns are achieved by preceding the regular character with an ESCape.

12 / 11 8 / 0 9 — The Hollerith card code for 256 characters is constructed from 1 or 2 or 3 . . . or 7 or blank (no punch) and any combination of 12, 11, 0, 8, and 9 (from none to all)

$$8 \times 32\ (2^5) = 256$$

For historical reasons, the assignments present little in the way of a regular pattern, but they are the key to translate to and from IBM EBCDIC.

REFERENCE CHART ISO CODE AND ASSOCIATED RELATIONSHIPS

Note — this is not a standard in itself. Refer to the appropriate documents (see reverse side). Screened characters in columns 3, 4, 5, 8 and 9 are under consideration.

Reprints of this chart are available from the Honeywell Computer Journal (P.O. Box 6000, Phoenix, AZ 85005) at $1 each postpaid.

col ' ıl (7-bit code)
te . ıar mnemonic
ne nic (8-bit code)
nn... onic (if different)

ıledge
ac·
ter lete
ter Insert
ter Insert Off
ter Insert On
ne ım Cursor
e F ırn
cre from Cursor
Control 1
Control 2
Control 3
Co ıl 4
ık ape
Medium
ne
Tra nission
ed ft In
ed Shift Out
Transmission Block
Tex
sec
par r
Sep ıtor
ıt Off
e Feed
e Reverse Feed
ht (
ıta b
ıta. b Clear
ıtal Tab Set
opy Off (full duplex)
opy On (half duplex)
let
ed
ıer
e Acknowledge
ıge
(Po er) Down
Fo at Off
Fo.. at On
(Pointer) Home
(Pointer) Left
(Pointer) Return
ıs F e
(Po er) Right
(Po er) Tab
(Pointer) Up
e Line Feed
Se ator
low
ıt
Heading
a blank)
Te
ıp
ute
onc... Idle
parator
l Tab
l Tab Clear
l Ta et

Note 1

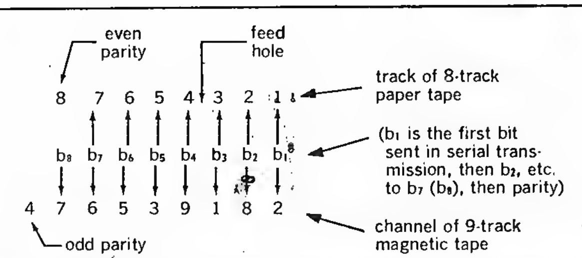
These 12 positions are variable — 2 for currency, 7 primary national usage, and 3 secondary usage which are diacritical marks used for alphabetical extension when preceded by BS. Positions 2/7 and 2/12 are invariant but also serve as diacritical marks. The presently-known assignments are given in the table below.

st 7 national			dia	dia	1st 7 national			dia
/1	5/12	5/13	5/14	6/0	7/11	7/12	7/13	7/14
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

(ESC) (CHAR.) SOFT COPY CONTI

$b_8\ b_7\ b_6\ b_5$ →			0000	0001	0010	0011	0100	0101	0110	[illegible]
	$b_4\ b_3\ b_2\ b_1$	ROW \ COL	0	1	2	3	4	5	6	
	0000	0	NUL □ 12-0-1-8-9	DLE 12-11-1-8-9	SP △	0 · 0	NOTE 1 @ CI 4-8	P CD 11-7	NOTE 1 ` 1-8	[illegible]
	0001	1	SOH 12-1-9	DC1 11-1-9	! 12-7-8	1 HTS 1	A PU 12-1	Q CIN 11-8	a 12-0-1	c
	0010	2	STX ⊥ 12-2-9	DC2 11-2-9	" 7-8	2 HTC 2	B PD 12-2	R CIF 11-9	b 12-0-2	r
	0011	3	ETX 12-3-9	DC3 11-3-9	NOTE 1 £/# 3-8	3 HN 3	C PR 12-3	S SU 0-2	c 12-0-3	s
	0100	4	EOT 7-9	DC4 4-8-9	NOTE 1 $/¤ 11-3-8	4 HF 4	D PL 12-4	T SD 0-3	d 12-0-4	t
	0101	5	ENQ ⊠ 0-5-8-9	NAK 5-8-9	% 0-4-8	5 VTS 5	E 12-5	U NP 0-4	e 12-0-5	u
	0110	6	ACK ✓ 0-6-8-9	SYN 2-9	& 12	6 VTC 6	F 12-6	V PP 0-5	f 12-0-6	v
	0111	7	BEL 0-7-8-9	ETB 0-6-9	' 5-8	7 RLF 7	G PM 12-7	W PFN 0-6	g 12-0-7	w
	1000	8	BS ↰ 11-6-9	CAN ⧖ 11-8-9	(12-5-8	8 HLR 8	H PH 12-8	X PF 0-7	h 12-0-8	x
	1001	9	HT → 12-5-9	EM 11-1-8-9	) 11-5-8	9 HLF 9	I PT 12-9	Y 0-8	i 12-0-9	y
	1010	10	LF ≡ 0-5-9	SUB ʕ 7-8-9	* 11-4-8	: LCF 2-8	J CSC 11-1	Z 0-9	j 12-11-1	z
	1011	11	VT ↓ 12-3-8-9	ESC ⊖ 0-7-9	+ 12-6-8	; LCN 11-6-8	K CLC 11-2	NOTE 1 [12-2-8	k 12-11-2	NOTE {
	1100	12	FF 12-4-8-9	FS 11-4-8-9	, 0-3-8	< 12-4-8	L LI 11-3	NOTE 1 \ 0-2-8	l 12-11-3	NOTE ¦/
	1101	13	CR ← 12-5-8-9	GS 11-5-8-9	− 11	= 6-8	M LD 11-4	NOTE 1] 11-2-8	m 12-11-4	NOTE }
	1110	14	SO ⊗ 12-6-8-9	RS 11-6-8-9	. 12-3-8	> 0-6-8	N 11-5	NOTE 1 ^ 11-7-8	n 12-11-5	~ NOTE
	1111	15	SI ⊙ 12-7-8-9	US 11-7-8-9	/ 0-1	? 0-7-8	O 11-6	_ 0-5-8	o 12-11-6	DE ▨

Alternate controls in these 5 columns are achieved by preceding the regular character with an ESCape.



REFERENCE C

Note — this is not a sta side). Screened charac

JISCII (Japanese Industrial Standard Code for Information Interchange) is an 8-bit code consisting of the ISO characters plus the Kata Kana characters shown in the upper row positions of columns 10-13 (columns 8 and 9 are reserved for additional controls, 14 and 15 for additional graphics).

GOST 13052-67 defines the USSR set shown in the lower row entry positions of columns 12-15. Actually, the standard defines these characters for columns 4-7 of a 7-bit set (SO = Russian register, SI = Latin register). Columns 8-11 are identical to 0-3.

(ESC) (CHAR.) SOFT COPY CONTROLS (SINGLE CHAR.)

	0100 4	0101 5	0110 6	0111 7	1000 8	1001 9	1010 10	1011 11	1100 12	1101 13	1110 14	1111 15
0	NOTE 1 @ / CI (4, 8)	P / CD (7, 11)	NOTE 1 ` (1, 8)	p (12 7, 11)	CI (1, 11 8, 0 9)	CD (12 1, 11 8, 0 9)	(12 1, 0 9)	ー (12, 11 8, 9)	タ / ю (12 6, 11, 0 9)	ミ / п (12 7, 11 8, 0)	Ю (12, 11 8, 0)	П (12 4, 11 8, 0 9)
1	A / PU (12 1)	Q / CIN (11 8)	a (12 1, 0)	q (12, 11 8)	PU (1, 0 9)	CIN (1, 9)	。 (12 2, 0 9)	ア (1, 11 8)	チ / а (12 7, 11, 0 9)	ム / я (1, 11 8, 0)	А (12, 11, 0 9)	Я (12 5, 11 8, 9)
2	B / PD (12 2)	R / CIF (11, 9)	b (12 2, 0)	r (12, 11, 9)	PD (2, 0 9)	CIF (2, 11 8, 9)	「 (12 3, 0 9)	イ (2, 11, 0 9)	ツ / б (12, 11 8, 0 9)	メ / р (2, 11 8, 0)	Б (12 2, 11 8, 0)	Р (12 6, 11 8, 9)
3	C / PR (12 3)	S / SU (2, 0)	c (12 3, 0)	s (2, 11, 0)	PR (3, 0 9)	SU (3, 9)	」 (12 4, 0 9)	ウ (3, 11, 0 9)	テ / ц (12 1, 8, 0)	モ / с (3, 11 8, 0)	Ц (12 3, 11 8, 0)	С (12 7, 11 8, 9)
4	D / PL (12 4)	T / SD (3, 0)	d (12 4, 0)	t (3, 11, 0)	PL (4, 0 9)	SD (4, 9)	、 (12 5, 0 9)	エ (4, 11, 0 9)	ト / д (12 2, 8, 0)	ヤ / т (4, 11 8, 0)	Д (12 4, 11 8, 0)	Т (2, 11 8, 0 9)
5	E (12 5)	U / NP (4, 0)	e (12 5, 0)	u (4, 11, 0)	(5, 11, 9)	NP (5, 9)	・ (12 6, 0 9)	オ (5, 11, 0 9)	ナ / е (12 3, 8, 0)	ユ / у (5, 11 8, 0)	Е (12 5, 11 8, 0)	У (3, 11 8, 0 9)
6	F (12 6)	V / PP (5, 0)	f (12 6, 0)	v (5, 11, 0)	(12 6, 9)	PP (6, 9)	ヲ (12 7, 0 9)	カ (6, 11, 0 9)	ニ / ф (12 4, 8, 0)	ヨ / ж (6, 11 8, 0)	Ф (12 6, 11 8, 0)	Ж (4, 11 8, 0 9)
7	G / PM (12 7)	W / PFN (6, 0)	g (12 7, 0)	w (6, 11, 0)	PM (7, 11, 9)	PFN (12, 8, 9)	ァ (12, 8, 0 9)	キ (7, 11, 0 9)	ヌ / г (12 5, 8, 0)	ラ / в (7, 11 8, 0)	Г (12 7, 11 8, 0)	В (5, 11 8, 0 9)
8	H / PH (12, 8)	X / PF (7, 0)	h (12, 8, 0)	x (7, 11, 0)	PH (8, 0 9)	PF (8, 9)	ィ (12 1, 8)	ク (11 8, 0 9)	ネ / х (12 6, 8, 0)	リ / ъ (12 1, 11 8, 0)	Х (12 2, 8, 0 9)	Ъ (6, 11 8, 0 9)
9	I / PT (12, 9)	Y (8, 0)	i (12, 0 9)	y (11 8, 0)	PT (1, 8, 0 9)	(1, 8, 9)	ゥ (12 1, 11, 9)	ケ (1, 8, 0)	ノ / и (12 7, 8, 0)	ル / ы (12 1, 11, 0)	И (12 3, 8, 0 9)	Ы (7, 11 8, 0 9)
10	J / CSC (1, 11)	Z (0 9)	j (12 1, 11)	z (11, 0 9)	CSC (2, 8, 0 9)	(2, 8, 9)	ェ (12 2, 11, 9)	コ (12, 11, 0)	ハ / й (12 1, 11 8)	レ / з (12 2, 11, 0)	Й (12 4, 8, 0 9)	З (12 2, 11 8, 0 9)
11	K / CLC (2, 11)	NOTE 1 [(12 2, 8)	k (12 2, 11)	NOTE 1 { (12, 0)	CLC (3, 8, 0 9)	(3, 8, 9)	ォ (12 3, 11, 9)	サ (12 1, 11, 0 9)	ヒ / к (12 2, 11 8)	ロ / ш (12 3, 11, 0)	К (12 5, 8, 0 9)	Ш (12 3, 11 8, 0 9)
12	L / LI (3, 11)	NOTE 1 \ (2, 8, 0)	l (12 3, 11)	NOTE 1 ¦/| (12, 11)	LI (4, 8, 0 9)	(12 4, 9)	ャ (12 4, 11, 9)	シ (12 2, 11, 0 9)	フ / л (12 3, 11 8)	ワ / э (12 4, 11, 0)	Л (12 6, 8, 0 9)	Э (12 4, 11 8, 0 9)
13	M / LD (4, 11)	NOTE 1] (2, 11 8)	m (12 4, 11)	NOTE 1 } (11, 0)	LD (12 1, 8, 9)	(4, 11, 9)	ュ (12 5, 11, 9)	ス (12 3, 11, 0 9)	ヘ / м (12 4, 11 8)	ン / щ (12 5, 11, 0)	М (12 7, 8, 0 9)	Щ (12 5, 11 8, 0 9)
14	N (5, 11)	NOTE 1 ^ (7, 11 8)	n (12 5, 11)	~/‾ NOTE 1 (1, 11, 0)	(12 7, 8, 9)	(6, 8, 9)	ョ (12 6, 11, 0)	セ (12 4, 11, 0 9)	ホ / н (12 5, 11 8)	゛ / ч (12 6, 11, 0)	Н (12 2, 11 8, 9)	Ч (12 6, 11 8, 0 9)
15	O (6, 11)	_ (5, 8, 0)	o (12 6, 11)	DEL (12 7, 9)	(3, 11 8, 9)	(1, 11, 0 9)	ッ (12 7, 11, 9)	ソ (12 5, 11, 0 9)	マ / о (12 6, 11 8)	゜ (12 7, 11, 0)	О (12 3, 11 8, 9)	EO (12 7, 11 8, 0 9)

Alternate controls in these 5 columns are achieved by preceding the regular character with an ESCape.

12	
11	8
0	9

The Hollerith card code for 256 characters is constructed from 1 or 2 or 3 . . . or 7 or blank (no punch) and any combination of 12, 11, 0, 8, and 9 (from none to all)

8
× 32 (2⁵)
= 256

For historical reasons, the assignments present little in the way of a regular pattern, but they are the key to translate to and from IBM EBCDIC.

REFERENCE CHART ISO CODE AND ASSOCIATED RELATIONSHIPS

Note — this is not a standard in itself. Refer to the appropriate documents (see reverse side). Screened characters in columns 3, 4, 5, 8 and 9 are under consideration.

Reprints of this chart are available from the Honeywell Computer Journal (P.O. Box 6000, Phoenix, AZ 85005) at $1 each postpaid.

bit
tra
b c.
n ty)
ack

5-char control (7-bit code)	alternate 2-char mnemonic	ANSI mnemonic (8-bit code)	ECMA mnemonic (if different)	
ACK	(AK)			Acknowledge
BEL	(BL)			Bell
BS				Backspace
CAN	(CN)			Cancel
		CD	DCL	Character Delete
		CI	ICL	Character Insert
		CIF		Character Insert Off
		CIN		Character Insert On
		CLC	EEL	Clear Line from Cursor
CR				Carriage Return
		CSC	EED	Clear Screen from Cursor
DC1	(D1)			Device Control 1
DC2	(D2)			Device Control 2
DC3	(D3)			Device Control 3
DC4	(D4)			Device Control 4
DEL	(DT)			Delete
DLE	(DL)			Data Link Escape
EM				End of Medium
ENQ	(EQ)			Enquiry
		EO		Eight Ones
EOT	(ET)			End of Transmission
ESC	(EC)			Escape
		ESI		Extended Shift In
		ESO		Extended Shift Out
ETB	(EB)			End of Transmission Block
ETX	(EX)			End of Text
FF				Form Feed
FS				File Separator
GS				Group Separator
		HF	RPM	Highlight Off
		HLF	FHL	Half Line Feed
		HLR	RHL	Half Line Reverse Feed
		HN	APM	Highlight On
HT				Horizontal Tab
		HTC		Horizontal Tab Clear
		HTS		Horizontal Tab Set
		LCF		Local Copy Off (full duplex)
		LCN		Local Copy On (half duplex)
		LD	DL	Line Delete
LF				Line Feed
		LI	IL	Line Insert
NAK	(NK)			Negative Acknowledge
		NP		Next Page
NUL	(NU)			Null
		PD	CUD	Cursor (Pointer) Down
		PFF		Protect Format Off
		PFN		Protect Format On
		PH	CUH	Cursor (Pointer) Home
		PL	CUB	Cursor (Pointer) Left
		PM		Cursor (Pointer) Return
		PP		Previous Page
		PR	CUF	Cursor (Pointer) Right
		PT		Cursor (Pointer) Tab
		PU	CUU	Cursor (Pointer) Up
		RLF		Reverse Line Feed
RS				Record Separator
		SD		Scroll Down
SI				Shift In
SO				Shift Out
SOH	(SH)			Start of Heading
SP				Space (a blank)
STX	(SX)			Start of Text
		SU		Scroll Up
SUB	(SB)			Substitute
SYN	(SY)			Synchronous Idle
US				Unit Separator
VT				Vertical Tab
		VTC		Vertical Tab Clear
		VTS		Vertical Tab Set

Note 1

These 12 positions are variable — 2 for currency, 7 ~~primary national usage, and 3~~ secondary usage which are diacritical marks used for alphabetical extension when preceded by BS. Positions 2/7 and 2/12 are invariant but also serve as diacritical marks. The presently-known assignments are given in the table below.

	currency		1st 7 national				dia	dia	1st 7 national			dia
	2/3	2/4	4/0	S/11	S/12	S/13	5/14	6/0	7/11	7/12	7/13	7/14
Netherlands—A			@	[illegible]	\	[illegible]	^	`	{	[illegible]	}	~
Australia	#											
Belgium—A												
W. Germany—A		$										
US												
Japan					¥							
UK	£				\		↑					
Italy—A	#	¤					^					
Switzerland—A		¤					^					
France—A		$										
USSR		¤			∨		∧					
Netherlands—B					[illegible]					[illegible]		
Belgium—B			[illegible]		ç				[illegible]	[illegible]	[illegible]	
France—B	£	$	à	°	ç	§	^	`	é	ù	è	[illegible]
Switzerland—B			[illegible]		ç				[illegible]	[illegible]	[illegible]	
Italy—B	#	$	§	°	ç	é	^	ù	à	ò	è	ì
Switzerland—C					ç	é		ü	[illegible]	[illegible]	[illegible]	[illegible]
Hungary	#	Ft	§	É	Ö	Ü	Á	ó	é	ő	ű	[illegible]
W. Germany—B	£	$	§	Ä	Ö	Ü	^	`	ä	ö	ü	ß
Switzerland—D				Ä	Ö	Ü			ä	ö	ü	ß
Sweden	#	¤	É	Ä	Ö	Å	Ü	é	ä	ö	å	ü
Finland				Ä	Ö	Å			ä	ö	å	[illegible]
Denmark				Æ	Ø	Å			æ	ø	å	[illegible]
Norway				Æ	Ø	Å			æ	ø	å	[illegible]
Spain					Ñ					ñ		

(ESC CHAR.) SOFT COPY CONTROLS (SINGLE CHAR.)

JISCII (Japanese Industrial Standard Code for Information Interchange) is an 8-bit code consisting of the ISO characters plus the Kata Kana characters shown in the upper-row positions of columns 10-13 (columns 8 and 9 are reserved for additional controls, 14 and 15 for additional graphics).

GOST-13052-67 defines the characters shown in the lower-row entry of columns 12-15. Actually defines these characters in 7-bit set with SO/SI (equivalent to Latin). Columns 8-11 are identical.

b_7 b_6 b_5 b_4 / b_4 b_3 b_2 b_1	ROW \ COL	0000 0	0001 1	0010 2	0011 3	0100 4	0101 5	0110 6	0111 7	1000 8	1001 9	1010 10	1011 11	1100 12	1101 13	1110 14
0000	0	NUL	DLE	SP	0	@ (NOTE 2) / CI	P / CD	` (NOTE 1)	p	CI	CD		ー	タ / ю	ミ / п	Ю
0001	1	SOH	DC1	!	1 / HTS	A / PU	Q / CIN	a	q	PU	CIN	。	ア	チ / а	ム / я	А
0010	2	STX	DC2	"	2 / HTC	B / PD	R / CIF	b	r	PD	CIF	「	イ	ツ / б	メ / р	Б
0011	3	ETX	DC3	£ / # (NOTE 1)	3 / HN	C / PR	S / SU	c	s	PR	SU	」	ウ	テ / ц	モ / с	Ц
0100	4	EOT	DC4	$ / ¤ (NOTE 1)	4 / HF	D / PL	T / SD	d	t	PL	SD	、	エ	ト / д	ヤ / т	Д
0101	5	ENQ	NAK	%	5 / VTS	E	U / NP	e	u		NP	・	オ	ナ / е	ユ / у	Е
0110	6	ACK	SYN	&	6 / VTC	F	V / PP	f	v		PP	ヲ	カ	ニ / ф	ヨ / ж	Ф
0111	7	BEL	ETB	'	7 / RLF	G / PM	W / PFN	g	w	PM	PFN	ァ	キ	ヌ / г	ラ / в	Г
1000	8	BS	CAN	(	8 / HLR	H / PH	X / PF	h	x	PH	PF	ィ	ク	ネ / х	リ / ъ	Х
1001	9	HT	EM	)	9 / HLF	I / PT	Y	i	y	PT		ゥ	ケ	ノ / и	ル / ы	И
1010	10	LF	SUB	*	: / LCF	J / CSC	Z	j	z	CSC		ェ	コ	ハ / й	レ / з	Й
1011	11	VT	ESC	+	; / LCN	K / CLC	[(NOTE 1)	k	{ (NOTE 1)	CLC		ォ	サ	ヒ / к	ロ / ш	К
1100	12	FF	FS	,	<	L / LI	\ (NOTE 1)	l	| (NOTE 1)	LI		ャ	シ	フ / л	ワ / э	Л
1101	13	CR	GS	−	=	M / LD	] (NOTE 1)	m	} (NOTE 1)	LD		ュ	ス	ヘ / м	ン / щ	М
1110	14	SO	RS	.	>	N	^ (NOTE 1)	n	‾ / ~ (NOTE 1)			ョ	セ	ホ / н	゛ / ч	Н
1111	15	SI	US	/	?	O	_	o	DEL			ッ	ソ	マ / о	゜	О

Alternate controls in these 5 columns are achieved by preceding the regular character with an ESCape.

12 / 11 8 / 0 9 — The Hollerith card code for 256 characters is constructed 1 or 2 or 3 . . . or 7 or blank (no punch) and any combination of 12, 11, 0, 8, and 9 (from none to all)

For historical reasons, the assignments present little in the way of a regular pattern, but they are the key to translate to and from IBM EBCDIC.

even parity — feed hole

8	7	6	5	4	3	2	1	← track of 8-track paper tape
b_8	b_7	b_6	b_5	b_4	b_3	b_2	b_1	← (b_1 is the first bit sent in serial transmission, then b_2, etc. to b_7 (b_8), then parity)
4	7	6	5	3	9	1	8	2 ← channel of 9-track magnetic tape

odd parity

REFERENCE CHART ISO CODE AND ASSOCIATED RELATIONSHIPS

Note — this is not a standard in itself. Refer to the appropriate documents (see reverse side). Screened characters in columns 3, 4, 5, 8 and 9 are under consideration.

Reprints of this chart are available from the Honeywell Computer Jo[urnal] (P.O. Box 6000, Phoenix, AZ 850[..]) at $1 each postpaid.

HONEYWELL BULL INTEROFFICE CORRESPONDENCE

RECEIVED

1973 JAN 2

ERIC H. CLAMONS

DATE: December 26, 1972

FROM: L. de Bournonville

LB/YL/1751/72

MAIL STATION: 2L053 E

TO: M. E. Clamons

MAIL STATION: Waltham - MS 453

COPIES TO: MM. Feissel
Mouren
Mac Namara (Waltham)

SUBJECT: Alternate Set

Please find attached a contribution, written by me in agreement with Mr. Feissel. I intend to present it at the next ECMA TC1 meeting (end of January).

This contribution has been sent to the members of the special interest list - codes.

Regards,

L. de Bournonville

Encl : 1

140

E C M A

EUROPEAN COMPUTER MANUFACTURERS ASSOCIATION

Alternate Set

by L. de Bournonville (Honeywell)

December 1972

In ECMA 6, some positions have alternate meanings or are reserved for National letters or National currency symbols.
In some applications, it is necessary to use more than one meaning for one position in the set.
This problem could be solved by the extension technique.
In the attached chart, we have gathered the alternate meanings already registered in national standards (and other standards).
This chart must be considered as a base for defining a standard alternate position using the extension technique, for the corresponding characters. It must be stated that :

1°) Some characters, such as quote and diaeresis may only be distinguishable by the font design of the character.
In the attached chart, we have given the OCR-B reference whenever the corresponding character was standard in OCR-B.

2°) This set goes against the rule of the uniqueness of character allocation as stated in ECMA 6.

With extension techniques, this chart could be used to solve problems such as :

- register of name and address with national letters
- incompatibility in National letter allocation in ECMA 6 and EBCDIC etc.

ALTERNATE SET

POSITION	MEANING	OCR-B REFER.
2-2	Quotation Mark	72
2-3	Pound Sign	89
2-4	Dollar Sign	90
2-5	Yen Sign	
2-7	Apostrophe	73
2-8	Russian Accent	
2-10	Ten Sign	
2-11	Eleven Sign	
2-12	Comma	69
2-13	Degree (Scandinavian accent)	
2-14	One Half	
3-2	Diaeresis	107
3-3	Number sign	84
3-4	Currency sign	88
3-5	Hungarian currency sign	
3-6	Cent sign	
3-7	Acute accent	108
3-8	Hacek	
3-10	Colon	70
3-11	Semi colon	71
3-12	Cedilla	112
3-13	Paragraph sign	
4-0	Commercial at	86
4-1	A with grave accent	
4-2	A with diaeresis	94
4-3	A with Scandinavian accent	95
4-4	E with grave accent	
4-6	N with Tilde	101
4-8	O with diaeresis	97
4-10	U with diaeresis (German U)	99
4-13	Dutch IJ	100
4-14	Upward arrow head	87
5-0	A with Tilde	
5-2	DIPHTHONG Æ	96
5-3	C with cedilla	
5-6	O with tilde	
5-8	O with slash	98
5-11	Left square bracket	81
5-12	Reverse Solidus	93
5-13	Right square bracket	82
5-14	Circumflex accent	110
6-0	Grave accent	109
6-1	a with acute accent	37 + 108
6-2	a with diaeresis	37 + 107
6-3	a with Scandinavian accent	102
6-4	e with acute accent	41 + 108
6-6	n with tilde	60 + 111
6-7	o with acute accent	51 + 108
6-8	o with diaeresis	51 + 107
6-10	u with diaeresis (German u)	57 + 107
6-12	Broken vertical line	

....

POSITION	MEANING	OCR-B REFER.
6-13	Small Dutch ij	105
6-14	Tilde	111
6-15	German double s	106
7-0	a with tilde	37 + 111
7-1	a with grave accent	37 + 109
7-2	Diphthong æ	103
7-3	c with cedilla	39 + 112
7-4	e with grave accent	41 + 109
7-5	i with grave accent	
7-6	o with tilde	51 + 111
7-7	o with grave accent	51 + 109
7-8	o with slash	104
7-10	u with grave accent	57 + 109
7-11	Left curly bracket	113
7-12	Vertical line	91
7-15	Right curly bracket	114

ALTERNATE SET

b_4	b_3	b_2	b_1	b_7 / b_6 / b_5	0 / 0 / 0	0 / 0 / 1	0 / 1 / 0	0 / 1 / 1	1 / 0 / 0	1 / 0 / 1	1 / 1 / 0	1 / 1 / 1
					0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	0	0	0	**0**					@	Ã	`	ã
0	0	0	1	**1**					Á		á	à
0	0	1	0	**2**			"	¨	Ä	Æ	ä	æ
0	0	1	1	**3**			£	#	Å	Ç	å	ç
0	1	0	0	**4**			$	¤	Ê		ê	è
0	1	0	1	**5**			¥	Ft				ì
0	1	1	0	**6**				Ç	Ñ	Õ	ñ	õ
0	1	1	1	**7**			'	´			ó	ò
1	0	0	0	**8**			ˇ	˘	Ö	Ø	ö	ø
1	0	0	1	**9**								
1	0	1	0	**10**			10	:	Ü		ü	ù
1	0	1	1	**11**			11	;		[		{
1	1	0	0	**12**			,	„		\	¦	|
1	1	0	1	**13**			°	§	IJ	]	ÿ	}
1	1	1	0	**14**			½		^	^	~	¯
1	1	1	1	**15**							ß	

TELEX

TO: P. MOUREN - XPR

FROM: E. H. CLAMONS, M/S 453 Waltham

DATE: 1/4/73

PLEASE REFERENCE L. DE BOURNONVILLE MEMO OF 72-12-26, ALTERNATE SETS.

SINCE HONEYWELL MANAGEMENT HAS NO FURTHER INTEREST IN CODE STANDARDIZATION, I SEE NO NEED IN PERPETUATING THE ACTIVITIES OF ECMA TC1. IN PARTICULAR, THE CONTRIBUTION OF M. DE BOURNONVILLE IS NOT MATCHED BY A CHB OR A HONEYWELL MARKETING EFFORT. I RECOMMEND THAT WE TAKE A PASSIVE ROLE IN ECMA TC1.

I WOULD APPRECIATE YOUR CHECKING THIS WITH M. FOUROT.

REGARDS,

ERIC H. CLAMONS

Honeywell Interoffice Correspondence

Date: January 9, 1973

To: L. de Bournonville

From: Eric H. Clamons

Location: Waltham, M/S 453

Subject: ALTERNATE SETS

I have reviewed your paper of 72/12/26, Alternate Sets. Here are my technical comments (personal, to you only).

1. The characters which you have put together for registration merely represent a summary of submissions to date from the various standards bodies. These will change as more and more countries register their codes. The registration you propose will change often and will be of little marketing value.

2. An additional page of characters would slow down the output speed of a printer, yet only a few characters would be used in each country. We would be well advised to limit ourselves to print trains which are specialized for each country and for each application.

3. The national problems are not solved entirely. For instance, only the more common accented French letters are included in your proposal; the rest are not necessary for typewritten material. However, if the customer pays a price for extra characters in France, he is usually not interested in the characters of Sweden, Norway, etc.

4. In certain countries; e.g., Eastern Europe, there is no limited set.

Therefore, from a marketing point of view, I would recommend that we develop special print trains, some for business correspondence - others for publishing, as needed. I recommend against developing an international standard which later could tie our hands competitively.

Eric H. Clamons

Eric H. Clamons

EHC:la

3 1

Table 19.17 Accented letters used in various languages

Afrikaans	à á â ä è é ê ë ì í î ï ò ó ô ö ù ú
Albanian	â ç ë
Czech	á č ď é ě í ň ó ř š ť ú ů ý ž Ď Ť
Danish	å ø
Dutch	è é ê ë ò ó ô ij
Estonian	ä ö õ š ü ž
Finnish	ä å ö
Flemish	ë ó ij
French	à â ç é è ê ë î ï ô ù û ü
Friesian	ä ê û
Gaelic (Scottish)	à é è ì ó ò ù
German	ä ö ü ß
Hungarian	á é í ó ö ő ú ü ű
Italian	à è é ì í î ò ó ù ú
Lithuanian	ą č ė ę š ū ų ž
Norwegian	å ø
Polish	ą ć ę ł ń ó ś ź ż
Portuguese	à á â ã ç è é ê ì í ò ó ô õ ú ù
Roumanian	à â ă è ì î ş ţ ù
Spanish	á é í ó ú ü ñ
Swedish	å ä ö

Table 19.17 lists the accented letters which are currently used in 21 languages which employ the latin alphabet. Additional accents may appear in the older literature of the countries using these languages and other accents could be required in modern texts where words have been borrowed or are quoted from another language. This subject is discussed in greater detail in *The Monotype Recorder Vol. 42 No. 4 Languages of the World.*

note {æ Æ} and {œ Œ} are part of "normal" character set

Table 19.16 Accented Letters for 11 European Languages

Roman	
Capitals	Á À Ä Â Å Ã É È Ë Ê Í Ì Ï Î Ó Ò Ö Ô Ø Õ Ú Ù Ü Û Ñ Ç
Lower case	á à ä â å ã é è ë ê í ì ï î ó ò ö ô ø õ ú ù ü û ñ ç ß ¿ ¡ « »
Small capitals	Á À Ä Â Å Ã É È Ë Ê Í Ì Ï Î Ó Ò Ö Ô Ø Õ Ú Ù Ü Û Ñ Ç
Italic	
Capitals	*Á À Ä Â Å Ã É È Ë Ê Í Ì Ï Î Ó Ò Ö Ô Ø Õ Ú Ù Ü Û Ñ Ç*
Lower case	*á à ä â å ã é è ë ê í ì ï î ó ò ö ô ø õ ú ù ü û ñ ç ß*
Bold	
Capitals	**Á À Ä Â Å Ã É È Ë Ê Í Ì Ï Î Ó Ò Ö Ô Ø Õ Ú Ù Ü Û Ñ Ç**
Lower case	**á à ä â å ã é è ë ê í ì ï î ó ò ö ô ø õ ú ù ü û ñ ç ß**

These accents, when added to a 7-alphabet fount will enable text composition in Danish, Dutch, English, French, Flemish, German, Italian, Spanish, Swedish, Norwegian and Portuguese

COMMITTEE CORRESPONDENCE

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

secretariat: CBEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466–2299

Doc. No. X3-73-90
Date July 30, 1973
Project 12
Milestone 9

Reply to:

RWB? OR who?

TO: Members of X3, SPARC, IAC, SSC, SP&PC, DPG/SC, X3S3
Officers of X3/TC's and SPARC/SG's
Members of X4, X4/TG on Keyboards, OMG/SC

SUBJECT: Draft Proposed Revision of ASCII, X3.4-1968

1. Attached for information and comment, if desired, is Doc. X3L2/1355, draft proposed revision of American National Standard Code for Information Interchange (ASCII).

2. This draft is currently being circulated for letter ballot within Technical Committee X3L2. The ballot closes September 7, 1973. All recipients are invited to submit comments to X3L2. For maximum usefullness, they should be received in time for consideration together with X3L2's consideration of its members' ballot comments, i.e., not later than September 7, 1973. Comments should be addressed to:

Mr. Charles D. Card (Chairman X3L2)
UNIVAC - Division of Sperry Rand Corp.
Site 2C1 - P. O. Box 500
Blue Bell, Pennsylvania 19422
(215) 542 3675

3. ACTION BY CHAIRMEN of SPARC, IAC, Technical Committees:

Each Chairman is requested to review this proposal for impact on his committee's work. If the chairman feels his committee will have substantive comments to make, he should immediately advise the Chairman of X3/SSC and Chairman of X3L2:

a. That his committee will probably make substantive comments,

b. The probable issues his committee will raise, and

c. The date by which SSC and X3L2 will receive his committee's comment.

In making such response, all Chairmen are urged to note the September 7th closing date of the X3L2 ballot. They should make every attempt to provide as much of their committees' views as possible by that date. SSC Section Managers are requested to monitor this review by their respective sections.

NM Att 3/0/7

Robert M. Brown
Robert M. Brown
Secretary, X3

Attachment: X3L2/1355

X3L2/1355
(X3L2/1304 revised)
1973 June 12

Project : 12
Milestone: 9

DRAFT PROPOSED REVISION
AMERICAN NATIONAL STANDARD
CODE FOR
INFORMATION INTERCHANGE

HOW TO KNOW:

1. CHNGS FROM PREVIOUS (& WHY? - §)
2. Δ FROM ISO R 646

This is a Working Document of Technical Committee X3L2, issued for review and comment. It is subject to revision or withdrawal.

National Standards Committee

X3

Computers and Information Processing

Secretariat:
Computer and Business Equipment Manufacturers Association

AMERICAN NATIONAL STANDARD CODE FOR INFORMATION INTERCHANGE

1. Scope

This coded character set is to be used for the general interchange of information among information processing systems, communications systems, and associated equipment.

2. Standard Code

				b_7	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1		0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	NL	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	§
1	1	0	1	13	CR	GS	–	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL

3. Character Representation and Code Identification

The standard 7-bit character representation, with b7 the high-order bit and b1 the low-order bit, is shown below:

Example: The bit representation for the character "K", positioned in column 4, row 11, is

b7	b6	b5	b4	b3	b2	b1
1	0	0	1	0	1	1

The code table for the character "K" may also be represented by the notation "column 4, row 11" or alternatively as "4/11." The decimal equivalent of the binary number formed by bit b7, b6, and b5, collectively, forms the column number, and the decimal equivalent of the binary number formed by bits b4, b3, b2, and b1, collectively, forms the row number.

The standard code may be identified by the use of the notation ASCII.

The notation ASCII (pronounced as'-key) should ordinarily be taken to mean the code prescribed by the latest issue of the standard. To explicitly designate a particular (perhaps prior) issue, the last two digits of the year of issue may be appended, as, "ASCII 68" or "ASCII 73".

NOT USASCII

4. Legend

4.1 Control Characters

Col/ Row	Mnemonic	Meaning	Col/ Row	Mnemonic	Meaning
0/0	NUL	Null	1/0	DLE	Data Link Escape (CC)
0/1	SOH	Start of Heading (CC)	1/1	DC1	Device Control 1
0/2	STX	Start of Text (CC)	1/2	DC2	Device Control 2
0/3	ETX	End of Text (CC)	1/3	DC3	Device Control 3
0/4	EOT	End of Trans- Mission (CC)	1/4	DC4	Device Control 4
0/5	ENQ	Enquiry (CC)	1/5	NAK	Negative Acknowledge (CC)
0/6	ACK	Acknowledge (CC)	1/6	SYN	Synchronous Idle (CC)
0/7	BEL	Bell	1/7	ETB	End of Transmission Block (CC)
0/8	BS	Backspace (FE)	1/8	CAN	Cancel
0/9	HT	Horizontal Tabulation (FE)	1/9	EM	End of Medium
0/10	NL	New Line (FE)	1/10	SUB	Substitute
0/11	VT	Vertical Tabulation (FE)	1/11	ESC	Escape
0/12	FF	Form Feed (FE)	1/12	FS	File Separator (IS)
0/13	HR	Horizontal Return (FE)	1/13	GS	Group Separator (IS)
0/14	SO	Shift Out	1/14	RS	Record Separator (IS)
0/15	SI	Shift in	1/15	US	Unit Separator (IS)
			7/15	DEL	Delete

NOTE: (CC) Communication Control
(FE) Format Effector
(IS) Information Separator

NOTE DROPPED

4.2 Graphic Characters

Column/Row	Symbol	Name
2/0	SP	Space (Normally Non-Printing)
2/1	!	Exclamation Point
2/2	"	Quotation Mark (Dieresis). See note 1.
2/3	# NOTE 4 £	Number Sign. See note 2.
2/4	$	Dollar Sign
2/5	%	Percent Sign
2/6	&	Ampersand
2/7	'	Apostrophe (Closing Single Quotation Mark; Acute Accent). See note 1.
2/8	(	Opening Parenthesis
2/9	)	Closing Parenthesis
2/10	*	Asterisk
2/11	+	Plus
2/12	,	Comma (Cedilla). See note 1.
2/13	-	Hyphen (Minus)
2/14	.	Period (Decimal Point)
2/15	/	Slant
3/0 to 3/9	0...9	Digits 0 to 9 inclusive
3/10	:	Colon
3/11	;	Semicolon
3/12	<	Less Than
3/13	=	Equals
3/14	>	Greater Than
3/15	?	Question Mark
4/0	@	Commercial At. See note 2.
4/1 to 5/10	A...Z	Uppercase Latin letters A thru Z
5/11	[	Opening Bracket. See note 2.
5/12	\	Reverse Slant. See note 2.
5/13	]	Closing Bracket. See note 2.
5/14	^	Circumflex. See notes 1 and 2.
5/15	_	Underline
6/0	`	Opening Single Quotation Mark
6/1 to 7/10	a...z	Lowercase Latin letters a thru z
7/11	{	Opening Brace. See note 2.
~~7/12~~	~~§~~	Section Mark. See note 2.
7/13	}	Closing Brace. See note 2.
7/14	~	Tilde (Overline). See notes 1 and 2.

1. The use of the symbols in 2/2, 2/7, 2/12, 5/14, 6/0, and 7/14 as diacritical marks is described in Appendix A, A5.2

2. These characters should not be used in international interchange without determining that there is agreement between sender and recipient. (See Appendix B4).

5. Definitions

5.1 General

ADDED

Control Character: A character whose occurrence in a particular context initiates, modifies, or stops an action that affects the recording, processing, transmission or interpretation of data.

Graphic Character: A character, other than a control character, that has a visual representation normally handwritten, printed or displayed.

(CC) Communication Control: A control character intended to control or facilitate transmission of information over communication networks.

(FE) Format Effector: A control character that controls the layout or positioning of information in printing or display devices.

(IS) Information Separator: A control character that is used to separate and qualify information in a logical sense. There is a group of four such characters, which are to be used in a hierarchical order.

5.2 Control Characters

4

0/0 NUL (Null): A control character used to accomplish media fill or time fill. Null characters may be inserted into or removed from a stream of data without affecting the information content of that stream. However, the addition or removal of these characters may affect the information layout or the control of equipment.

3

0/1 SOH (Start of Heading): A communication control character used as the first character of a heading of an information message.

0/2 STX (Start of Text): A communication control character that precedes a text and is used to terminate a heading.

0/3 ETX (End of Text): A communication control character that terminates a text.

0/4 EOT (End of Transmission): A communcation control character used to indicate the conclusion of a transmission, which may have contained one or more texts and any associated headings.

0/5 ENQ (Enquiry): A communication control character used in data communication systems as a request for a response from a remote station. It may be used as a "Who Are YOU" (WRU) to obtain identification, or may be used to obtain station status, or both.

0/6 ACK (Acknowledge): A communication control character transmitted by a receiver as an affirmative response to a sender.

0/7 BEL (Bell): A control character for use when there is a need to call for attention. It may control alarm or attention devices.

0/8 BS (backspace): A one-position format effector that moves the active position backwards on the same line.

0/9 HT (Horizontal Tabulation): A format effector that advances the active position to the next predetermined character position on the same line.

0/10 NL (New Line): A format effector that advances the active position to the first active position on the next line. Where appropriate, this character may have the meaning "Line Feed" (LF), a format effector which advances the active position to the same character position of the next line. Use of the LF convention requires agreement between sender and recipient of data.

WHY? (INDEX ON IBM TYPER)

0/11 VT (Vertical Tabulation): A format effector that advances the active position to the first character position on the next predetermined line. When agreed upon between the interchange parties VT may advance the active position to the same character position on the next predetermined line.

0/12 FF (Form Feed): A format effector that advances the active position to the first character position on a predetermined line of the next form or page. When agreed upon between the interchange parties, FF may advance the active position to the same character position on a predetermined line of the next form or page.

0/13 HR (Horizontal Return): A format effector that moves the active position to the first character position on the same line.

0/14 SO (Shift Out): A control character that is used in conjunction with Shift In to extend the graphic character set. It may alter the meaning of the bit combinations of columns 2 and 7 which follow it until a Shift In character is reached. However, the characters Space (2/0) and Delete (7/15) are unaffected. The effect of this character is described in the code extension standard (see section 7.0).

0/15 SI (Shift In): A control character that is used in conjunction with Shift Out to extend the graphic set. It may reinstate the standard meanings of the bit combinations which follow it. The effect of this character is described in the code extension standard (see section 7.0) 2

1/0 DLE (Data Link Escape): A communication control character that will change the meaning of a limited number of contiguously following characters. It is used exclusively to provide supplementary data transmission control functions. Appropriate sequences are defined in another standard (see section 7.0). 3

1/1 to 1/4 DC1, DC2, DC3, DC4 (Decice Controls): Control characters for the control of ancillary devices associated with data processing or telecommunication systems, more especially switching devices "on" or "off". (If a single "stop" control is required to interrupt or turn off ancillary devices, DC4 is the preferred assignment.)

1/5 NAK (Negative Acknowledge): A communication control character transmitted by a receiver as a negative response to the sender.

1/6 SYN (Synchronous Idle): A communication control character used by a synchronous transmission system in the absence of any other character to provide a signal from which synchronism may be achieved or retained.

1/7 ETB (End of Transmission Block): A communication control character used to indicate the end of a block of data for communication purposes. ETB is used for blocking data where the block structure is not necessarily related to the processing format.

1/8 CAN (Cancel): A control character used to indicate that the data with which it is sent is in error or is to be disregarded. The specific meaning of this character must be defined for each application.

1/9 EM (End of Medium): A control character that may be used to identify the physical end of a medium, or the end of the used portion of a medium, or the end of the wanted portion of data recorded on a medium. The position of this character does not necessarily correspond to the physical end of the medium.

1/10 SUB (Substitute): A character that may be substituted for a character that is determined to be invalid or in error.

1/11 ESC (Escape): A control character intended to provide supplementary characters (code extension). The Escape character itself is a prefix affecting the interpretation of a limited number of contiguous bit patterns. The effect of 2 this character is described in the code extension standard (see section 7.0).

1/12 to 1/15 FS (File Separator), GS (Group Separator), RS (Record Separator), and US (Unit Separator): These information separators may be used with data in optional fashion, except that their hierarchical relationship shall be: FS as the most inclusive, then GS, then RS, and US as least inclusive. (The content and length of a File, Group, Record, or Unit are not specified.)

7/15 DEL (Delete): A character used primarily to erase or obliterate an erroneous or unwanted character in punched tape. DEL characters may also serve to accomplish media-fill or time-fill. They may be inserted into or removed from a stream of data without affecting the information content of that stream. However, the addition or removal of these characters may affect the information layout and/ or the control of equipment.

TIME FILL + MEDIA FILL?

5.3 Graphic Characters

No specific meaning is prescribed for any of the graphics in the code table except that which is understood by the users. Furthermore, this standard does not specify a type style for the printing or display of the various graphic characters. In specific applications, it may be desirable to employ distinctive styling of individual graphics to facilitate their use for specific purposes.

2/0 SP (Space): A graphic character that is usually represented by a blank site in a series of graphics. The space character, though not a control character, has a function equivalent to that of a format effector that causes the active position to move one position forward without producing the printing or display of any graphic. Similarly, the space character may have a function equivalent to that of an information separator.

LOWEST ORDER?

"CLARIFIED"?

6. General Considerations

6.1 This standard does not define the means by which the coded set is to be recorded in any physical medium, nor does it include any redundancy or define techniques for error control. Further, this standard does not define data communication character structure, data communication formats, code extension techniques, or graphic representation of control characters.

6.2 Deviations from the standard may create serious difficulties in information interchange and should be used only with full cognizance of the parties involved.

7

6.3 The relative sequence of any two characters, when used as a basis for collation, is defined by their binary values. The Null character (position 0/0) will be ranked low and the Delete character (position 7/15) will be ranked high. Other collating sequences may be used by prior agreement between interchanging parties.

NEW

6.4 The representation of this 7-bit code in an 8-bit environment is specified in the standard on code extension (see section 7.0).

6.5 The appendixes to this standard contain additional information on the design and use of this code.

?

OLD 6.4 8 - GRAPHIC STYLIZATION

7. Related Standards

[illegible] 1199	Proposed American National Standard X3.41. Code Extension Techniques for Use With the 7-Bit Coded Character Set of ASCII.
IS 2022	Code Extension Techniques for Use With the 7-Bit Coded Character Set of IS 646.
IS 646	7-Bit Coded Character Set for Information Processing Interchange.
IS 1745	Basic Mode Control Procedures for Data Communications.
X3.6 - 1965	Perforated Tape Code for Information Interchange.
X3.16 - 1966	Character Structure and Character Parity Sense for Serial-by-Bit Data Communication in the American National Standard Code for Information Interchange.
X3.22 - 1967	Recorded Magnetic Tape for Information Interchange (800 CPI, NRZI).
X3.25 - 1968	Character Structure and Character Parity Sense for Parallel-by-Bit Data Communication in the American National Standard Code for Information Interchange.
X3.26 - 1970	Hollerith Punched Card Code.
X3.28 - 1971	Procedures for the Use of the Communication Control Characters of American National Standard Code for Information Interchange in Specified Data Communication Links.

ADDED

APPENDIXES

These appendixes are not a part of American National Standard Code for Information Interchange, X3.4-1973, but are included to facilitate its use.

APPENDIX A

DESIGN CONSIDERATIONS CHARACTER SET

A1. Introduction

The standard coded character set is intended for the interchange of information among information processing systems, communication systems, and associated equipment.

A2. Considerations Affecting the Code

There were many considerations that determined the set size, set structure, character selection, and character placement of the code. Among these were (not listed in order of priority):

(1) Need for adequate number of graphic symbols
(2) Need for adequate number of device controls, format effectors, communication controls, and information separators
(3) Desire for a nonambiguous code; i.e., one in which every code combination has a unique interpretation
(4) Physical requirements of media and facilities
(5) Error control considerations
(6) Special interpretation of the all-zeros and all-ones characters
(7) Ease in the identification of classes of characters
(8) Data manipulation requirements
(9) Collating conventions
 (a) Logical
 (b) Historical
(10) Keyboard conventions
 (a) Logical
 (b) Historical
(11) Other set sizes
(12) International considerations
(13) Programming languages
(14) Existing coded character sets.

A3. Set Size

A 7-bit set is the minimum size that will meet the requirements for graphics and controls in applications involving general information interchange.

A4. Set Structure

A4.1 In discussing the set structure it is convenient to divide the set into 8 columns and 16 rows as indicated in this standard.

A4.2 It was considered essential to have a dense subset which contained only graphics. For ease of identification this graphic subset was placed in six contiguous columns.

A4.3 The first two columns were chosen for the controls for the following reasons:

(1) The character NUL by its definition has the location 0/0 in the code table. NUL is broadly considered a control character.

(2) The representations in column 7 were felt to be most susceptible to simulation by a particular class of transmission error --- one which occurs during an idle condition on asynchronous systems.

(3) To permit the considerations of graphic subset structure described in A6 to be satisfied, the two columns of controls had to be adjacent.

A4.4 The character set was structured to enable the easy identification of classes of graphics and controls.

A5. Choice of Graphics

A5.1 Included in the set are the digits 0 through 9, upper and lower cases of the Latin letters A through Z, and those punctuation, mathematical, and business symbols considered most useful. The set includes a number of characters commonly encountered in programming languages. In particular, all the COBOL and FORTRAN graphics are included.

A5.2 In order to permit the representation of languages other than English, two diacritical (or accent) marks have been included, and provision has been made for the use of four punctuation symbols alternately as diacritical marks. The pairing of these punctuation symbols with their corresponding diacritical marks was done to facilitate the design of a typeface which would be acceptable for both uses.

These arrangements are:

Col/ Row	Code Table Symbol	Use: Punctuation	Use: Diacritical
2/2	"	Quotation Marks	Dieresis
2/7	'	Apostrophe	Acute Accent
2/12	,	Comma	Cedilla
5/14	^	(None)	Circumflex
6/0	`	Opening Single Quotation Mark	Grave Accent
7/14	~	(None)	Tilde

A6. Graphic Subset Structure

A6.1 The basic structure of the dense graphic subset was influenced by logical collating considerations, the requirements of simply related 6-bit sets, and the needs of typewriter-like devices. For information processing, it is desirable that the characters be arranged in such a way as to minimize both the operating time and the hardware components required for ordering and sequencing operations. This requires that the relative order of characters, within classes, be such that a simple comparison of the binary codes will result in information being ordered in a desired sequence.

A6.2 Conventional usage requires that SP (space) be ahead of any other symbol in a collatable set. This permits a name such as "JOHNS" to collate ahead of a name such as "JOHNSON." The requirement that punctuation symbols such as comma also collate ahead of the alphabet ("JOHNS, A" should also collate before "JOHNSON") establishes the special symbol locations including SP, in the first column of the graphic subset.

A6.3 To simplify the design of typewriter-like devices, it is desirable that there be only a common 1-bit difference between characters to be paired on keytops. This, together with the requirements for a contiguous alphabet, and the collating requirements outlined above, resulted in the placement of the alphabet in the last four columns of the graphic subset and the placement of the digits in the second column of the graphic subset.

A6.4 It is expected that devices having the capability of printing only 64 graphic symbols will continue to be important. It may be desirable to arrange these devices to print one symbol for the bit pattern of both upper and lower case of a given alphabetic letter. To facilitate this, there should be a single-bit difference between the upper and lower case representations of any given letter. Combined with the requirement that a given case of the alphabet be contiguous, this dictated the assignment of the alphabet, as shown in columns 4 through 7.

A6.5 To minimize ambiguity caused by the use of a 64-graphic device as described above, it is desirable to the degree possible, that each character in column 6 or 7 differ little in significance from the corresponding character in column 4 or 5. In certain cases, this was not possible.

6

A6.6 The assignment of reverse slant and section mark, the brackets and braces, and the circumflex and tilde were made with a view to the considerations of A6.5.

A6.7 The resultant structure of "specials" (S), "digits" (D), and "alphabetics" (A) does not conform to the most prevalant collating convention (S-A-D) because of other code requirements.

A6.8 The need for a simple transformation from the set sequence to the prevalent collating convention was recognized, and dictated the placement of some of the "specials" within the set. Specifically, these special symbols, viz., ampersand (&), asterisk (*), comma (,), hyphen (-), period (.), and slant (/), which are most often used as identifiers for ordering information and which normally collate ahead of both the alphabet and the digits were not placed in the column containing the digits, so that the entire numeric column could be rotated via a relatively simple transformation to a position higher than the alphabet. The sequence of the aforementioned "specials" was also established to the extent practical to conform to the prevalent collating convention.

A6.9 The need for a useful 4-bit numeric subset also played a role in the placement of characters. Such a 4-bit subset, including the digits and the symbols asterisk, plus (+), comma, hyphen, period, and slant, can easily be derived from the code.

A6.10 Considerations of other domestic code sets, including the Department of Defense former standard 8-bit data transmission code ("Fieldata"-based), as well as international requirements, played an important role in deliberation that resulted in the code. The selection and grouping of the symbols dollar sign ($), percent sign (%), ampersand (&), apostrophe ('), less than (<), equals (=), and greater than (>) facilitate contraction to either a business or scientific 6-bit subset. The position of these symbols, and of the symbols comma, hyphen, period, and slant, facilitates achievement of commonly accepted pairing on a keyboard. The historic pairing of question mark and slant is preserved and the less than and greater than symbols, which have comparatively low usage, are paired with period and comma, so that in dual-case keyboard devices where it is desired to have period and comma in both cases, the less than and greater than symbols are the ones displaced. Provision was made for the accommodation of alphabets containing more than 26 letters and for 6-bit contraction by the location of low-usage characters in the area following the alphabet. — DROPPED STERLING

A7. Control Subset Content and Structure

A7.1 The control characters included in the set are those required for the control of terminal devices, input and output devices, format, or communication transmission and switching, and are general enough to justify inclusion in a standard set.

A7.2 Many control characters may be considered to fall into the following categories:

(1) Communication Controls
(2) Format Effectors
(3) Device Controls
(4) Information Separators

To the extent practical, controls of each category were grouped in the code table. DROPPED MENTION OF RELATED 6-BIT SET

A7.3 The information separators (FS, GS, RS, US) identify boundaries of various elements of information, but differ from punctuation in that they are primarily intended to be machine sensible. They were arranged in accordance with an expected hierarchical use, and the lower end of the hierarchy is contiguous in binary order with SP (space) which is sometimes used as a machine-sensible separator. Subject to this hierarchy the exact nature of their use within data is not specified.

A7.4 The character SYN (Synchronous Idle) was located so that its binary pattern in serial transmission was unambiguous as to

character framing, and also to optimize certain other aspects of its communication usage.

A7.5 ACK (Acknowledge) and NAK (Negative Acknowledge) were located so as to gain the maximum practical protection against mutation of one into the other by transmission errors.

A7.6 The function "New Line" (NL) has been assigned as the principal control function to position 0/10 and "Line Feed" (LF) assigned as the secondary function. This has been incorporated in the standard to reflect the trend to a single function to perform the new line function. These assignments allow the use of a common end-of-line format for both printers having separate HR-LF functions and those having a combined (i.e., NL) function. The end-of-line format HR-LF will produce the same result on printers of both classes and would be useful during conversion of a system from one method of operation to the other. 1

A8. Collating Sequence

A8.1 This supplements the consideration of collating sequence in section A6.

> It is recognized that the collating sequence defined in paragraph 6.3 of this standard cannot be used in many specific applications that define their own sequence. In some applications, groups of characters may be assigned exactly equal collating weight to preserve an initial ordering. In other applications, a different sequence may be desired to meet the needs of the particular application. Nonetheless, it was deemed essential to define a standard sequence and standard results for comparisons of two items of data, to serve the needs of many applications. The standard sequence will facilitate, but will not provide directly by means of simple sorting, the ordering of items customarily found in (1) algebraically signed data, in which the largest positive value is high and the largest negative value is low (2) complex alphabetic listings, such as those found in telephone directories, library catalogs or dictionaries. However, general use of these standard collating sequences and standard comparison evaluations will facilitate the transfer of programs and the general interchange of data among various computer systems. 7

APPENDIX B

NOTES ON APPLICATION

B1. Introduction

B1.1 The standard code was developed to provide for information interchange among information processing systems, communications systems, and associated equipment. In a system consisting of equipment with several local or native codes, maximum flexibility will be achieved if each of the native codes is translated to the standard whenever information interchange is desired.

B1.2 Within any particular equipment, system, or community of activity, it may be necessary to substitute characters. For example, some systems may require special graphic symbols and some devices may require special control codes. (Design efforts on the standard code included consideration of these types of adaptations.) So-called "secular sets" produced by such substitutions, although not conforming to this standard, may nonetheless be consonant with it if substitutions are made in accordance with the guidelines of B2.

B2. Character Substitutions

B2.1 Any character substitution will result in a coded character set which does not conform to this standard.

B2.2 It is recommended that when a character is substituted in the code table for a standard character, the standard character should not be reassigned elsewhere in the table. Such a substitute character, once assigned, should not be subsequently reassigned elsewhere.

B2.3 It is recommended that graphic substitutions be made only in the graphic area and control substitutions only in the control area. Any substitution involving a control should be made only with full cognizance of all possible operational effects.

B2.4 It should be noted that this standard specifies, for each position of the code table, the information represented by the character and not necessarily the precise action taken by the recipient when the character is received. In the case of graphics, considerable variation in the actual shape printed or displayed may be appropriate to different units, systems, or fields of application. In the case of controls, the action performed is dependent upon the use for which the particular system is intended, the application to which it is being put, and

a number of conventions established by the user or designer --- some system-wide and some unique to a particular unit.

B2.5 Typical examples of diversity in execution not necessarily contrary to this standard are:

(1) A number of graphic symbols, other than that used in the code table, are used for <u>ampersand</u> in various type styles; still other symbols may be more appropriate to electronic display devices. The use of such alternate symbols does not in itself constitute deviation from the standard as long as <u>ampersand</u> is the concept associated with the character. Note that this does not necessarily restrict the use of such an alternate symbol design to mean "and"; in any type style <u>ampersand</u> may, of course, be used with arbitrary meaning.

(2) A card punch in one application may "skip" when the character HT (Horizontal Tabulation: used as skip) is presented to it; in another application the character HT may be recorded in the card without further action.

B3. Related Larger and Smaller Sets

Consideration has been given to the relationship between the standard set and sets of other sizes. A number of straightforward logical transformations are possible which result in a variety of sets related to the standard. None of the transformed sets are recognized by this standard.

B4. International Considerations

This standard conforms to the recommendations of the International Organization for Standardization (ISO) and the International Telegraph and Telephone Consultative Committee (CCITT)* for a 7-bit code. It includes all the character assignments specified by those bodies for international standardization. Their recommendations, however, permit national standardization by the various countries in seven code table positions. Also, the characters in three additional positions have been designated as being replaceable by national characters in only those countries having an extraordinary requirement in this regard.

The seven National Use positions, as well as the three "supplementary" National Use positions (which are denoted by a section mark), and their assignments in this standard are shown in the following:

Column/Row	Character (U.S.)
4/0	@
5/11	[
5/12	\
5/13	]
§5/14	^
§6/0	`
7/11	{
7/12	§
7/13	}
§7/14	~

In international interchange of information these 10 characters should not be used except where it is determined that there is agreement between sender and recipient.

In addition, in other countries, the number sign (#) (in position 2/3) may be replaced by "£", and the dollar sign ($) (in position 2/4) may be replaced by the currency symbol (¤).

---

*An international body which establishes standards and conventions for international telecommunications, especially where the public telegraph and telephone services are governmentally owned and operated. Their recommendations are often embodied in the regulations applying to such services.

B5. Communications Considerations

Certain control characters are designated as "communication controls." They are:

SOH	(Start of Heading)
STX	(Start of Text)
ETX	(End of Text)
EOT	(End of Transmission)
ENQ	(Enquiry)
ACK	(Acknowledge)
DLE	(Data Link Escape)
NAK	(Negative Acknowledge)
SYN	(Synchronous Idle)
ETB	(End of Transmission Block)

These may be used by communication systems for their internal signaling, or for the exchange of information relating to the control of the communication system between that system and its end terminals. Some such systems may impose restrictions on the use of these communication control characters by the end terminals. For example, the use of some of them may be completely prohibited while others may be restricted to use in conformity with the formats and procedures required by the communication system for its operation.

APPENDIX C

Original Criteria

C1.1 Introduction

C1.1 This Appendix contains the original criteria upon which the design of the code was based. Not all criteria have been entirely satisfied. Some are conflicting, and the characteristics of the set represent accepted compromises of these divergent criteria.

C1.2 The criteria were drawn from communication, processing, and media recording aspects of information interchange.

C2. Criteria

C2.1 Every character of the code set shall be represented by the same number of bits.

C2.2 The standard set shall be so structured as to facilitate derivation of logically related larger or smaller sets.

C2.3 In a code of n bits, all possible 2n patterns of ones and zeros will be permitted and considered valid.

C2.4 The number of bits, n, shall be sufficient to provide for the alphabitic and numeric characters, commonly used punctuation marks, and other special symbols, along with those control characters required for interchange of information.

C2.5 The digits 0 through 9 shall be included within a 4-bit subset.

C2.6 The digits 0 through 9 shall be so represented that the four low-order bits shall be the binary-coded-decimal form of the particular digit that the code represents. DROPPED "STORMING"

C2.7 The interspersion of control characters among the graphic characters shall be avoided. The characters devoted to controls shall be easily separable from those devoted to graphics.

C2.8 Within the standard set, each character shall stand by itself and not depend on surrounding characters for interpretation.

C2.9 An entire case of the Latin alphabet (A through Z) shall be included within a 5-bit subset. Consideration shall be given to the need for more than 26 characters in some alphabets.

C2.10 The letters of each case of the alphabet shall be assigned, in conventional order (A through Z), to successive, increasing binary representations.

C2.11 Suitable control characters required for communication and information processing shall be included.

C2.12 Escape functions that provide for departures from the standard set shall be incorporated.

C2.13 A simple binary comparison shall be sufficient to determine the order within each class of characters. (In this regard, the special graphics, the digits and the alphabet are each defined as distinct classes.) Simple binary rules do not necessarily apply between classes when ordering information.

C2.14 Space must collate ahead of all other graphics.

C2.15 Special symbols used in the ordering of information must collate ahead of both the alphabet and the digits.

C2.16 Insofar as possible, the special symbols shall be grouped according to their functions; for example, punctuation and mathematical symbols. Further, the set shall be so organized that the simplest possible test shall be adequate to distinguish and identify the basic alphabetic, numeric, and special symbol subsets.

C2.17 Special symbols shall be placed in the set so as to simplify their generation by typewriters and similar keyboard devices. This criterion means, in effect, that the codes for pairs of characters that normally appear on the same keytops on a typewriter shall differ only in a common single-bit position.

C2.18 The set shall contain the graphic characters of the principal programming languages.

C2.19 The codes for all control characters shall contain a common, easily recognizable, bit pattern.

C2.20 The Null (000...) and Delete (111...) characters shall be provided.

APPENDIX D

Terminology

This Appendix is intended to clarify the sense in which certain terms are used.

Active Position: That character position in which the character about to be processed would appear. The active position normally advances one character position at a time. ADDED

Bit: Contraction of "binary digit."

Bit Pattern: The binary representation of a character.

Character: A member of a coded character set; the binary representation of such a member and its graphic symbol or control function.

Code: A system of discrete representation of a set of symbols and functions.

INTERNATIONAL STANDARD ISO 646

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

7-bit coded character set for information processing interchange

First edition — 1973-07-01

UDC 681.3.042 : 003.62

Ref. No. ISO 646-1973 (E)

Descriptors : data processing, data transmission, character sets, coding, information interchange.

Price based on 12 pages

ISO 646-1973 (E)

FOREWORD

ISO (the International Organization for Standardization) is a worldwide federation of national standards institutes (ISO Member Bodies). The work of developing International Standards is carried out through ISO Technical Committees. Every Member Body interested in a subject for which a Technical Committee has been set up has the right to be represented on that Committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the Technical Committees are circulated to the Member Bodies for approval before their acceptance as International Standards by the ISO Council.

International Standard ISO 646 was drawn up by Technical Committee ISO/TC 97, *Computers and information processing*, and circulated to the Member Bodies in April 1972.

It has been approved by the Member Bodies of the following countries :

Belgium
Brazil
Canada
Czechoslovakia
Denmark
Egypt, Arab Rep. of
France
Germany
Ireland
Italy
Japan
Netherlands
Portugal
Romania
South Africa, Rep. of
Spain
Sweden
Switzerland
Thailand
United Kingdom
U.S.A.
U.S.S.R.

No Member Body expressed disapproval of the document.

Printed in Switzerland

7-bit coded character set for information processing interchange

1 SCOPE AND FIELD OF APPLICATION

1.1 This International Standard contains a set of 128 characters (control characters and graphic characters such as letters, digits and symbols) with their coded representation. Most of these characters are mandatory and unchangeable, but provision is made for some flexibility to accommodate special national and other requirements.

1.2 The need for graphics and controls in data processing and in data transmission has been taken into account in determining this character set.

1.3 This International Standard consists of a general table with a number of options, notes, a legend and explanatory notes. It also contains a specific international reference version, guidance on the exercise of the options to define specific national versions and application oriented versions.

1.4 This character set is primarily intended for the interchange of information among data processing systems and associated equipment, and within message transmission systems.

1.5 This character set is applicable to all latin alphabets.

1.6 This character set includes facilities for extension where its 128 characters are insufficient for particular applications.

1.7 The definitions of some control characters in this International Standard assume that data associated with them is to be processed serially in a forward direction. Their effect when included in strings of data which are processed other than serially in a forward direction or included in data formatted for fixed record processing may have undesirable effects or may require additional special treatment to ensure that the control characters have their desired effect.

2 IMPLEMENTATION

2.1 This character set should be regarded as a basic alphabet in an abstract sense. Its practical use requires definitions of its implementation in various media. For example, this could include punched tapes, punched cards, magnetic tapes and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate recording in a physical medium, or by local electrical connection of various units (such as input and output devices and computers) or by means of data transmission equipment.

2.2 The implementation of this coded character set in physical media and for transmission, taking into account the need for error checking, is the subject of other ISO publications. (See Appendix Y.)

3 BASIC CODE TABLE

TABLE 1 — Basic code table

				b_7	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1	column / row	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	TC_7 (DLE)	SP	0	[3]	P	` [4]	p
0	0	0	1	1	TC_1 (SOH)	DC_1	!	1	A	Q	a	q
0	0	1	0	2	TC_2 (STX)	DC_2	" [6]	2	B	R	b	r
0	0	1	1	3	TC_3 (ETX)	DC_3	£(#) [2]	3	C	S	c	s
0	1	0	0	4	TC_4 (EOT)	DC_4	$(¤) [2]	4	D	T	d	t
0	1	0	1	5	TC_5 (ENQ)	TC_8 (NAK)	%	5	E	U	e	u
0	1	1	0	6	TC_6 (ACK)	TC_9 (SYN)	&	6	F	V	f	v
0	1	1	1	7	BEL	TC_{10} (ETB)	' [6]	7	G	W	g	w
1	0	0	0	8	FE_0 (BS)	CAN	(	8	H	X	h	x
1	0	0	1	9	FE_1 (HT)	EM	)	9	I	Y	i	y
1	0	1	0	10	FE_2 (LF) [1]	SUB	*	:	J	Z	j	z
1	0	1	1	11	FE_3 (VT) [1]	ESC	+	;	K	[3]	k	[3]
1	1	0	0	12	FE_4 (FF) [1]	IS_4 (FS)	, [6]	<	L	[3]	l	[3]
1	1	0	1	13	FE_5 (CR) [1]	IS_3 (GS)	-	=	M	[3]	m	[3]
1	1	1	0	14	SO	IS_2 (RS)	.	>	N	^ [4][5]	n	‾ [4][5]
1	1	1	1	15	SI	IS_1 (US)	/	?	O	_	o	DEL

NOTES ABOUT TABLE 1

① The format effectors are intended for equipment in which horizontal and vertical movements are effected separately. If equipment requires the action of CARRIAGE RETURN to be combined with a vertical movement, the format effector for that vertical movement may be used to effect the combined movement. For example, if NEW LINE (symbol NL, equivalent to CR + LF) is required, FE_2 shall be used to represent it. This substitution requires agreement between the sender and the recipient of the data.

The use of these combined functions may be restricted for international transmission on general switched telecommunication networks (telegraph and telephone networks).

② The symbol £ is assigned to position 2/3 and the symbol $ is assigned to position 2/4. In a situation where there is no requirement for the symbol £ the symbol # (number sign) may be used in position 2/3. Where there is no requirement for the symbol $ the symbol ¤ (currency sign) may be used in position 2/4. The chosen allocations of symbols to these positions for international information interchange shall be agreed between the interested parties. It should be noted that, unless otherwise agreed between sender and recipient, the symbols £, $ or ¤ do not designate the currency of a specific country.

③ National use positions. The allocations of characters to these positions lies within the responsibility of national standardization bodies. These positions are primarily intended for alphabet extensions. If they are not required for that purpose, they may be used for symbols.

④ Positions 5/14, 6/0 and 7/14 are provided for the symbols UPWARD ARROW HEAD, GRAVE ACCENT and OVERLINE. However, these positions may be used for other graphical characters when it is necessary to have 8, 9 or 10 positions for national use.

⑤ Position 7/14 is used for the graphic character ‾ (OVERLINE), the graphical representation of which may vary according to national use to represent ~(TILDE) or another diacritical sign provided that there is no risk of confusion with another graphic character included in the table.

⑥ The graphic characters in positions 2/2, 2/7, 2/12 and 5/14 have respectively the significance of QUOTATION MARK, APOSTROPHE, COMMA and UPWARD ARROW HEAD; however, these characters take on the significance of the diacritical signs DIAERESIS, ACUTE ACCENT, CEDILLA and CIRCUMFLEX ACCENT when they are preceded or followed by the BACKSPACE character (0/8).

4 LEGEND

4.1 Control characters

Abbrevi-ation	Note	Meaning	Position in the code table
ACK		Acknowledge	0/6
BEL		Bell	0/7
BS		Backspace	0/8
CAN		Cancel	1/8
CR	1	Carriage Return	0/13
DC		Device control	–
DEL		Delete	7/15
DLE		Data Link Escape	1/0
EM		End of Medium	1/9
ENQ		Enquiry	0/5
EOT		End of Transmission	0/4
ESC		Escape	1/11
ETB		End of Transmission Block	1/7
ETX		End of Text	0/3
FE		Format Effector	–
FF	1	Form Feed	0/12
FS		File Separator	1/12
GS		Group Separator	1/13
HT		Horizontal Tabulation	0/9
IS		Information Separator	–
LF	1	Line Feed	0/10
NAK		Negative Acknowledge	1/5
NUL		Null	0/0
RS		Record Separator	1/14
SI		Shift-In	0/15
SO		Shift-Out	0/14
SOH		Start of Heading	0/1
SP		Space (see 7.2)	2/0
STX		Start of Text	0/2
SUB		Substitute Character	1/10
SYN		Synchronous Idle	1/6
TC		Transmission Control	–
US		Unit Separator	1/15
VT	1	Vertical Tabulation	0/11

4.2 Graphic characters

Graphic	Note	Name	Position in the code table
(space)		Space (see 7.2)	2/0
!		Exclamation mark	2/1
"	6	Quotation mark, Diaeresis	2/2
£	2	Pound sign	2/3
#	2	Number sign	2/3
$	2	Dollar sign	2/4
¤	2	Currency sign	2/4
%		Percent sign	2/5
&		Ampersand	2/6
'	6	Apostrophe, acute accent	2/7
(		Left parenthesis	2/8
)		Right parenthesis	2/9
*		Asterisk	2/10
+		Plus sign	2/11
,	6	Comma, Cedilla	2/12
–		Hyphen, Minus sign	2/13
.		Full stop (period)	2/14
/		Solidus	2/15
:		Colon	3/10
;		Semi-colon	3/11
<		Less than sign	3/12
=		Equals sign	3/13
>		Greater than sign	3/14
?		Question mark	3/15
^	4, 6	Upward arrow head, Circumflex accent	5/14
_		Underline	5/15
`	4	Grave accent	6/0
‾	4, 5	Overline, Tilde	7/14

5 EXPLANATORY NOTES

5.1 Numbering of the positions in Table 1

Within any one character the bits are identified by b_7, $b_6 \ldots b_1$, where b_7 is the highest order, or most significant bit, and b_1 is the lowest order, or least significant bit.

If desired, these may be given a numerical significance in the binary system, thus :

Bit identification :	b_7	b_6	b_5	b_4	b_3	b_2	b_1
Significance :	64	32	16	8	4	2	1

In the table the columns and rows are identified by numbers written in binary and decimal notations.

Any one position in the table may be identified either by its bit pattern, or by its column and row numbers. For instance, the position containing the digit 1 may be identified :

- by its bit pattern in order of decreasing significance, i.e. 011 0001;
- by its column and row numbers, i.e. 3/1.

The column number is derived from bits b_7, b_6 and b_5 giving them weights of 4, 2 and 1 respectively. The row number is derived from bits b_4, b_3, b_2 and b_1 giving them weights of 8, 4, 2 and 1 respectively.

5.2 Diacritical signs

In the 7-bit character set, some printing symbols may be designed to permit their use for the composition of accented letters when necessary for general interchange of information. A sequence of three characters, comprising a letter, "backspace" and one of these symbols, is needed for this composition; the symbol is then regarded as a diacritical sign. It should be noted that these symbols take on their diacritical significance when they are preceded or followed by one "backspace" character; for example the symbol corresponding to the code combination 2/7 normally has the significance of "apostrophe", but becomes the diacritical sign "acute accent" when it precedes or follows a "backspace" character.

In order to increase efficiency, it is possible to introduce accented letters (as single characters) in the positions marked by Note ③ in the code table. According to national requirements, these positions may contain special diacritical signs.

5.3 Names, meanings and fonts of graphic characters

This International Standard assigns at least one name to denote each of the graphic characters displayed in Tables 1 and 2. The names chosen to denote graphic characters are intended to reflect their customary meanings. However, this International Standard does not define and does not restrict the meanings of graphic characters. In addition, it does not specify a particular style or font design for the graphic characters.

Under the provision of Note ③ of Table 1, graphic characters which are different from the characters of the international reference version may be assigned to the national use positions. When such assignments are made, the graphic characters 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 international reference version.

5.4 Uniqueness of character allocation

A character allocated to a position in Table 1 may not be placed elsewhere in the table. For example, in the case of position 2/3 the character not used cannot be placed elsewhere. In particular the POUND sign (£) can never be represented by the bit combination of position 2/4.

6 VERSIONS OF TABLE 1

6.1 General

6.1.1 In order to use Table 1 for information interchange, it is necessary to exercise the options left open, i.e. those affected by Notes ② to ⑤. A single character must be allocated to each of the positions for which this freedom exists or it must be declared to be unused. A code table completed in this way is called a "version".

6.1.2 The notes to Table 1, the explanatory notes and the legend apply in full to any version.

6.2 International reference version

This version is available for use when there is no requirement to use a national or an application-oriented version. In international information processing interchange the international reference version (Table 2) is assumed unless a particular agreement exists between sender and recipient of the data. The following characters are allocated to the optional positions of Table 1 :

#	Number sign	2/3
¤	Currency sign	2/4
@	Commercial at	4/0
[	Left square bracket	5/11
\	Reverse solidus	5/12
]	Right square bracket	5/13
{	Left curly bracket	7/11
|	Vertical line	7/12
}	Right curly bracket	7/13

It should be noted that no substitution is allowed when using the international reference version.

TABLE 2 — International reference version

				b_7	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1	row \ column	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	TC_7 (DLE)	SP	0	@	P	`	p
0	0	0	1	1	TC_1 (SOH)	DC_1	!	1	A	Q	a	q
0	0	1	0	2	TC_2 (STX)	DC_2	"	2	B	R	b	r
0	0	1	1	3	TC_3 (ETX)	DC_3	#	3	C	S	c	s
0	1	0	0	4	TC_4 (EOT)	DC_4	¤	4	D	T	d	t
0	1	0	1	5	TC_5 (ENQ)	TC_8 (NAK)	%	5	E	U	e	u
0	1	1	0	6	TC_6 (ACK)	TC_9 (SYN)	&	6	F	V	f	v
0	1	1	1	7	BEL	TC_{10} (ETB)	'	7	G	W	g	w
1	0	0	0	8	FE_0 (BS)	CAN	(	8	H	X	h	x
1	0	0	1	9	FE_1 (HT)	EM	)	9	I	Y	i	y
1	0	1	0	10	FE_2 (LF)	SUB	*	:	J	Z	j	z
1	0	1	1	11	FE_3 (VT)	ESC	+	;	K	[	k	{
1	1	0	0	12	FE_4 (FF)	IS_4 (FS)	,	<	L	\	l	|
1	1	0	1	13	FE_5 (CR)	IS_3 (GS)	-	=	M	]	m	}
1	1	1	0	14	SO	IS_2 (RS)	.	>	N	^	n	‾
1	1	1	1	15	SI	IS_1 (US)	/	?	O	_	o	DEL

6.3 National versions

6.3.1 The responsibility for defining national versions lies with the national standardization bodies. These bodies shall exercise the options available and make the required selection.

6.3.2 If so required, more than one national version can be defined within a country. The different versions shall be separately identified. In particular when for a given national use position, for example, 5/12 or 6/0, alternative characters are required, two different versions shall be identified, even if they differ only by this single character.

6.3.3 If there is in a country no special demand for specific characters, it is strongly recommended that the characters of the international reference version be allocated to the same national use positions.

6.4 Application-oriented versions

Within national or international industries, organizations or professional groups, application-oriented versions can be used. They require precise agreement among the interested parties, who will have to exercise the options available and to make the required selection.

7 FUNCTIONAL CHARACTERISTICS RELATED TO CONTROL CHARACTERS

Some definitions in this section are stated in general terms and more explicit definitions of use may be needed for specific implementation of the code table on recording media or on transmission channels. These more explicit definitions and the use of these characters are the subject of other ISO Publications.

7.1 General designations of control characters

The general designation of control characters involves a specific class name followed by a subscript number.

They are defined as follows :

TC Transmission control characters

Control characters intended to control or facilitate transmission of information over telecommunication networks.

The use of the TC characters on the general telecommunication networks is the subject of other ISO publications.

The transmission control characters are :

ACK, DLE, ENQ, EOT, ETB, ETX, NAK, SOH, STX and SYN.

FE Format effectors

Control characters mainly intended for the control of the layout and positioning of information on printing and/or display devices. In the definitions of specific format effectors, any reference to printing devices should be interpreted as including display devices.

The definitions of format effectors use the following concept :

a) a page is composed of a number of lines of characters;

b) the characters forming a line occupy a number of positions called character positions;

c) the active position is that character position in which the character about to be processed would appear, if it were to be printed. The active position normally advances one character position at a time.

The format effector characters are : BS, CR, FF, HT, LF and VT (see also Note 1).

DC Device control characters

Control characters for the control of a local or remote ancillary device (or devices) connected to a data processing and/or telecommunication system. These control characters are not intended to control telecommunication systems; this should be achieved by the use of TCs.

Certain preferred uses of the individual DCs are given in 7.2.

IS Information separators

Control characters that are used to separate and qualify data logically. There are four such characters. They may be used either in hierarchical order or non-hierarchically; in the latter case their specific meanings depend on their applications.

When they are used hierarchically, the ascending order is :

US, RS, GS, FS.

In this case, data normally delimited by a particular separator cannot be split by a higher order separator but will be considered as delimited by any higher order separator.

7.2 Specific control characters

Individual members of the classes of controls are sometimes referred to by their abbreviated class name and a subscript number (for example, TC_5) and sometimes by a specific name indicative of their use (for example, ENQ).

Different but related meanings may be associated with some of the control characters but in an interchange of data this normally requires agreement between the sender and the recipient.

ACK Acknowledge

A transmission control character transmitted by a receiver as an affirmative response to the sender.

BEL Bell

A control character that is used when there is a need to call for attention; it may control alarm or attention devices.

BS Backspace

A format effector which moves the active position one character position backwards on the same line.

CAN Cancel

A character, or the first character of a sequence, indicating that the data preceding it is in error. As a result, this data is to be ignored. The specific meaning of this character must be defined for each application and/or between sender and recipient.

CR Carriage return

A format effector which moves the active position to the first character position on the same line.

Device controls

DC_1 A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to restore a device to the basic mode of operation (see also DC_2 and DC_3), or for any other device control function not provided by other DCs.

DC_2 A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to set a device to a special mode of operation (in which case DC_1 is used to restore the device to the basic mode), or for any other device control function not provided by other DCs.

DC_3 A device control character which is primarily intended for turning off or stopping an ancillary device. This function may be a secondary level stop, for example wait, pause, stand-by or halt (in which case DC_1 is used to restore normal operation). If it is not required for this purpose, it may be used for any other ancillary device control function not provided by other DCs.

DC_4 A device control character which is primarily intended for turning off, stopping or interrupting an ancillary device. If it is not required for this purpose, it may be used for any other device control function not provided by other DCs.

Examples of use of the device controls :

1) One switching		
	on–DC_2	off–DC_4
2) Two independent switchings		
first one	on–DC_2	off–DC_4
second one	on–DC_1	off–DC_3
3) Two dependent switchings		
general	on–DC_2	off–DC_4
particular	on–DC_1	off–DC_3
4) Input and output switching		
output	on–DC_2	off–DC_4
input	on–DC_1	off–DC_3

DEL Delete

A character used primarily to erase or obliterate an erroneous or unwanted character in punched tape. DEL characters may also serve to accomplish media-fill or time-fill. They may be inserted into or removed from a stream of data without affecting the information content of that stream but then the addition or removal of these characters may affect the information layout and/or the control of equipment.

DLE Data link escape

A transmission control character which will change the meaning of a limited number of contiguously following characters. It is used exclusively to provide supplementary data transmission control functions. Only graphic characters and transmission control characters can be used in DLE sequences.

EM End of medium

A control character that may be used to identify the physical end of a medium, or the end of the used portion of a medium, or the end of the wanted portion of data recorded on a medium. The position of this character does not necessarily correspond to the physical end of the medium.

ENQ Enquiry

A transmission control character used as a request for a response from a remote station – the response may include station identification and/or station status. When a "Who are you" function is required on the general switched transmission network, the first use of ENQ after the connection is established shall have the meaning "Who are you" (station identification). Subsequent use of ENQ may, or may not, include the function "Who are you", as determined by agreement.

EOT End of transmission

A transmission control character used to indicate the conclusion of the transmission of one or more texts.

ESC Escape

A control character which is used to provide additional control functions. It alters the meaning of a limited number of contiguously following bit combinations. The use of this character is specified in ISO 2022.

ETB End of transmission block

A transmission control character used to indicate the end of a transmission block of data where data is divided into such blocks for transmission purposes.

ETX End of text

A transmission control character which terminates a text.

FF Form feed

A format effector which advances the active position to the same character position on a pre-determined line of the next form or page.

HT Horizontal tabulation

A format effector which advances the active position to the next pre-determined character position on the same line.

Information separators

IS_1 (US) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a UNIT.

IS_2 (RS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a RECORD.

IS_3 (GS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a GROUP.

IS_4 (FS) A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a FILE.

LF Line feed

A format effector which advances the active position to the same character position of the next line.

NAK Negative acknowledge

A transmission control character transmitted by a receiver as a negative response to the sender.

NUL Null

A control character used to accomplish media-fill or time-fill. NUL characters may be inserted into or removed from a stream of data without affecting the information content of that stream; but then the addition or removal of these characters may affect the information layout and/or the control of equipment.

SI Shift-in

A control character which is used in conjunction with SHIFT-OUT and ESCAPE to extend the graphic character set of the code. It may reinstate the standard meanings of the bit combinations which follow it. The effect of this character when using code extension techniques is described in ISO 2022.

SO Shift-out

A control character which is used in conjunction with SHIFT-IN and ESCAPE to extend the graphic character set of the code. It may alter the meaning of the bit combinations of columns 2 to 7 which follow it until a SHIFT-IN character is reached. However, the characters SPACE (2/0) and DELETE (7/15) are unaffected by SHIFT-OUT. The effect of this character when using code extension techniques is described in ISO 2022.

SOH Start of heading

A transmission control character used as the first character of a heading of an information message.

SP Space

A character which advances the active position one character position on the same line.

This character is also regarded as a non-printing graphic.

STX Start of text

A transmission control character which precedes a text and which is used to terminate a heading.

SUB Substitute character

A control character used in the place of a character that has been found to be invalid or in error. SUB is intended to be introduced by automatic means.

SYN Synchronous idle

A transmission control character used by a synchronous transmission system in the absence of any other character (idle condition) to provide a signal from which synchronism may be achieved or retained between data terminal equipment.

VT Vertical tabulation

A format effector which advances the active position to the same character position on the next pre-determined line.

APPENDIX Y

RELEVANT ISO PUBLICATIONS

ISO/R 961, *Implementation of the 6 and 7-bit coded character sets on 7 track 12,7 mm (1/2 in) magnetic tape.*

ISO/R 962, *Implementation of the 7-bit coded character set on 9 track 12,7 mm (1/2 in) magnetic tape.*

ISO/R 963, *Guide for the definition of 4-bit character sets derived from the ISO 7-bit coded character set for information processing interchange.*

ISO/R 1113, *Representation of 6 and 7-bit coded character sets on punched tape.*

ISO/R 1155, *The use of longitudinal parity to detect errors in information messages.*

ISO/R 1177, *Character structure for start/stop and synchronous transmission.*

ISO/R 1679, *Representation of ISO 7-bit coded character set on 12-row punched cards.*

ISO/R 1745, *Basic mode control procedures for data communication systems.*

ISO 2022, *Code extension techniques for use with the ISO 7-bit coded character set.*

ISO 2047, *Graphical representations for the control characters of the ISO 7-bit coded character set.*

APPENDIX Z

ISO Recommendation R 646-1967 contained a table for a 6-bit coded character set. This character set has been deleted from the present International Standard and is no more part of the standard. The corresponding table with the relevant Notes is reproduced in this Appendix for information only.

Bits				b_6 →	0	0	1	1
				b_5 →	0	1	0	1
b_4	b_3	b_2	b_1	Row \ Column	0	1	2	3
0	0	0	0	0	SP	0	NUL	P
0	0	0	1	1	F_1 (HT)	1	A	Q
0	0	1	0	2	F_2 (LF) ①	2	B	R
0	0	1	1	3	F_3 (VT)	3	C	S
0	1	0	0	4	F_4 (FF)	4	D	T
0	1	0	1	5	F_5 (CR) ①	5	E	U
0	1	1	0	6	SO	6	F	V
0	1	1	1	7	SI	7	G	W
1	0	0	0	8	(	8	H	X
1	0	0	1	9	)	9	I	Y
1	0	1	0	10	*	: ⑧	J	Z
1	0	1	1	11	+	; ⑧	K	([) ③
1	1	0	0	12	,	< ⑨ $ ②	L	(£) ② ③
1	1	0	1	13	—	= ⑨ %	M	(]) ③
1	1	1	0	14	.	> ⑨ &	N	ESC
1	1	1	1	15	/	'	O	DEL

① The controls CR and LF are intended for printer equipment which requires separate combinations to return the cariage and to feed a line.

For equipment which uses a single control for a combined cariage return and line feed operation,

- in the 6-bit set table, the function F_2 will have the meaning of "New Line" (NL), F_5 will then have the meaning of "Back-space" (BS);
- in the 7-bit set table, the function FE_2 will have the meaning of "New Line" (NL).

These substitutions require agreement between the sender and the recipeint of the data.

The use of this function "NL" is not allowed for international transmission on general telecommunication networks (Telex and Telephone networks).

② For international information interchange, $ and £ symbols do not designate the currency of a given country. The use of these symbols combined with other graphic symbols to designate national currencies may be the subject of other ISO Recommendations.

③ Reserved for National Use. These positions are primarily intended for alphabetic extensions. If they are not required for that purpose, they may be used for symbols and a recommended choice is shown in parenthesis in some cases.

Some restrictions are placed on the use of these characters on the general telecommunication networks for international transmission.

⑧ If 10 and 11 as single characters are needed (for example, for Sterling currency subdivision), they should take the place of "colon" (:) and "semi-colon" (;) respectively. These substitutions require agreement between the sender and the recipient of the data.

On the general telecommunication networks, the characters "colon" and "semi-colon" are the only ones authorized for international transmission.

⑨ Either of the two sets of three symbols shown in these positions in the table may be chosen; the interpretation of the corresponding combinations requires agreement between the sender and the recipient of the data.

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

Doc. No. ADM
Date 9/26/73
Project

COMMITTEE CORRESPONDENCE

secretariat: BEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466-2288

Reply to: Charles D. Card
Sperry Univac
P.O. Box 500
Blue Bell, PA 19422
M. S. 2C1

MEMO TO: X3L2 Members, Alternates, Consultants, and Observers

FROM: C. D. Card, Chairman X3L2

SUBJECT: Enclosed documents

Enclosed are two X3L2 documents (X3L2/1389 and 1390). Both of these documents have attachments. Please make sure that you have these attachments.

Charles D Card
C. D. Card (MBR)

CDC:mbr

Attachments

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

secretariat: BEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466-2288

Doc. No. X3L2/1389
Date 9/24/73
Project 5D

COMMITTEE CORRESPONDENCE

Reply to: Charles D. Card
Sperry Univac
P.O. Box 500
Blue Bell, PA 19422
M. S. 2C1

TO: Edward M. Piehl--CBEMA

FROM: Charles D. Card, Chairman X3L2

CC: X3L2 Members, Alternates, Consultants, and Observers

SUBJECT: X4 Draft Combined Keyboard

Your document of 10 September arrived too late for our committee to review at its meeting 11-13 September.

The following comments are from the writer as an individual and do not represent the committee.

X3L2, of course, has no official position on the following:

E00	± ⑀	
E03	£	
E04	¤	
E06	¢	
E11	⑁	
D11	1/4 1/2	
D12	¦ § ¶	
B08	■	(group erase)
B09	—	(character erase)

The reason is that those symbols are not in ASCII and we have had no request from X4, X3A1, or others to add them in our "Graphics For an 8-Bit Code" project.

The problem we envision is that, when added (or registered), they may have 8-bit patterns with the high order bit equal one. Thus the positioning becomes moot.

My own position is that the OCR "controls" should be optional graphics for the control functions they represent. Along the same lines I prefer control X for group erase.

C D Card

C. D. Card

CDC:mbr

american national standards committees:
X3–Computers & Information Processing
X4–Office Machines & Supplies

secretariat: CBEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466–2299

COMMITTEE CORRESPONDENCE

Doc. No.
Date September 10, 1973
Project
Milestone

Reply to:

TO: X4, X4/Keyboards, X3L2, X3A1, X3

SUBJECT: Working Draft on Combined Keyboard

Attached is a keyboard layout for typewriters developed by X4/Keyboards at its August 21-22 meeting. A draft document is currently being prepared that will be based upon this layout. The layout shows the placement of the characters of X4.7, OCR-A, OCR-B (97 character subset) and ASCII.

The characters are tied to the keys on which they are placed and cannot be moved (except in the case of ʃ which may be placed above period and comma when key B00 is deopped and upper case period and comma are not needed).

The placement of OCR-B characters group erase (B09) and character erase (B08) are shown, although this placement is not definite because of possible problems (which are being investigated; this placement is the one being considered by X4/Keyboards at the present time.

X3L2 and X3A1 are requested to review the attached layout and forward any comments to X4/Keyboards by November 1, 1973 to my attention at CBEMA, address above.

Edd
Edward M. Piehl
Assistant Director
Standards

SEP 17 1973
CUSTOMER SUPPORT
FACILITY – 1048

/snd

Attachment

DRAFT KEYBOARD LAYOUT COMBINING THE CHARACTER SETS OF OCR-A, OCR-B, ASCII AND X4.7

	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	
F																		F
E			~±ʃ ˙°ʜ	! 1	@ 2	#£ 3	$¤ 4	% 5	^¢^ 6	& 7	* 8	(9	) 0	Ч _ -	+ =			E
D				Q	W	E	R	T	Y	U	I	O	P	] ¾ [½	¦ § | \ ¶ \			D
C				A	S	D	F	G	H	J	K	L	: ;	" '	} ³ { ²			C
B			> <	Z	X	C	V	B	N	M	,<▮ ,	.> — .	? /					B
A								SPACE BAR										A
Z																		Z

Approved as a Working Group Draft
Date 1973 August 22

NOTE: The location of OCR Characters, GROUP ERASE and CHARACTER ERASE (B08 and B09), are not fully approved. Their final placement will be deferred until typewriter manufacturers provide input on the effect of these two characters in these locations.

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

Doc. No. X3L2/1390
Date 9/24/73
Project 12M

COMMITTEE CORRESPONDENCE

secretariat: BEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466-2288

Reply to: Charles D. Card
Sperry Univac
P.O. Box 500
Blue Bell, PA 19422
M. S. 2C1

MEMO TO: Kent Godwin and Michael Marcotty

CC: R.E. Schubert, X3L2 Members, Alternates, Consultants, and Observers

FROM: C. D. Card, Chairman X3L2

SUBJECT: Response to concerns on revision of X3.4- 1968.

The technical sub-committee X3L2 has taken into account your communications. The weight of Mr. Godwin's argument was felt in accord with the desires of the group. As a result, the next revision to be balloted will contain a solid vertical line at position 7/12 in the table. It will be called Vertical Line as in the past.

The attention of X3J1 is drawn to this matter. X3L2 found itself in a potential cross fire and opted toward the direction of compatibility with the existing standards.

CDCa

C. D. Card
Chairman X3L2

CDC:mbr

Attachment

COMMITTEE CORRESPONDENCE

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

secretariat: BEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466-2299

Doc. No. SSC/317

Date 1973 August 23

Project 12

Milestone

Reply to: G. Kent Godwin
7817 Glenister Drive
Springfield, Va. 22152

Mr. Robert M. Brown
Chairman, SSC
CBEMA
1828 L Street, N.W.
Washington, D.C. 20036

Dear Bob:

I have just read document X3-73-90, Draft Proposed Revision to ASCII, X3.4-1968.

This proposal directly impacts X3A1 Projects 57, 59, and 61 through the change of the graphic symbol and name of the character in position 7/12.

X3.4-1968 contained the character Vertical Line. This revision contains the character Section Mark.

Project 57, Revision of X3.17-1966 contains the character vertical line.

Project 59, OCR-ASCII extension contains the character vertical line.

Project 61, OCR-B (BSR X3.49) contains the character Vertical Line.

In addition, ISO R1073 and IS 2033 both contain the character Vertical Line.

The introduction of a replacement graphic symbol at this late date, unaccompanied by any explanation, creates a difficult situation.

On behalf of X3A1, I request that X3L2 reconsider the proposed change in position 7/12 and revert to the Vertical Line.

Sincerely,

G. Kent Godwin

G. Kent Godwin, Manager
Recognition, X3/SSC

cc:
C.D. Card
R.E. Schubert

computing

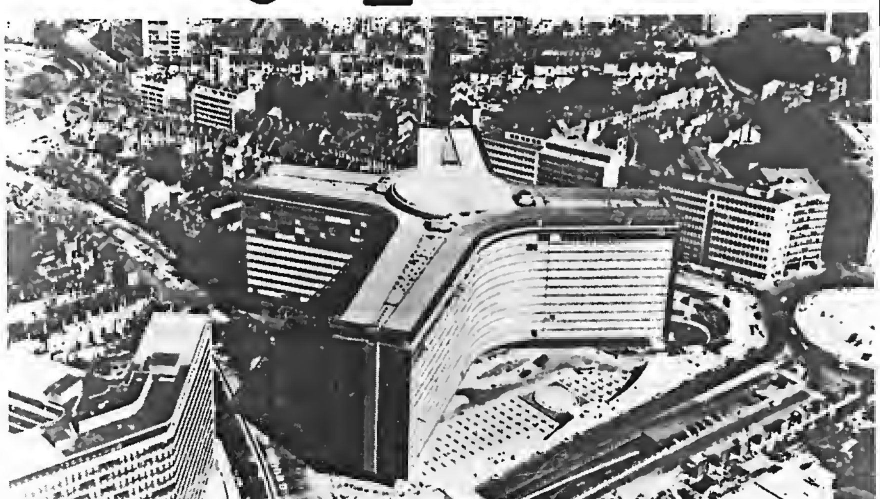
Published weekly by Haymarket Publishing Ltd for The British Computer Society 29/11/73

NMAH 310/7

EEC releases its industry plans

A procedure for the joint planning and management of data networks on a European basis is one of the proposals in a report of the European Commission presented to the Council of Ministers in Brussels last week.

For the consolidation of the European computer industry, the report says urgent priority should be given to the development of an effective application of common standards in hardware and software. One possible form of support would be a European leasing company to back the industry, it suggests.



User group contradicts Grosch theory

A rapid reaction to the suggestion in last week's *Computing* interview with Herb Grosch that the IBM user groups SHARE and GUIDE were getting weaker came from a GUIDE spokesman this week. He claimed that the European trend in this area was in the opposite direction.

The spokesman said that the last GUIDE meeting was attended by some 3,500 members compared with the average 2,000 plus.

At this week's Mini Guide meeting at Heathrow, near London Airport, it was revealed that another IBM users' group, SEAS, is working together with GUIDE and a third group, COMMON, on long term plans for computer requirements over a time scale of up to 10 years hence.

IMS was among the subjects discussed at the meeting.

Bob Bemer

Bemer and the nine bit byte

Bob Bemer latched on to the idea of the nine bit byte by piecing together several items of information. Chairman of the International Standards Organisation's committee on programming languages, Bemer now works for Honeywell after a career spanning several manufacturers. Nancy Foy writes about the nine lives of Bemer.

Profile page 8.

DOS/VS and Topsy

(a task which had never before been done mechanically). 'Then a round little man approached and asked what I

lull of 1963 when people believed that an era of co-operation and goodwill

IBM could use a nine-bit byte on the new line, which was to be announced 75-wise. With this inspiration, I talked it up.

questions in Europe, where national computer companies would have to spend millions to make similar

users: We found the technology to give you everything, *and* standards, without destroying any little part of

software system that IBM ever supplied on pre-announced schedule,' adds Bemer-the-Honeyweller.

Bemer-the-professional

Bemer's nine lives

Bob Bemer is one of those types of computer pioneer who inevitably turns up just about where the action is next year, particularly if the action has anything to do with the standards we work towards or the social effects of our professional decisions. He has been with IBM, and Univac and General Electric USA and thus with Honeywell. When Nancy Foy asked him what he did for a living Bemer replied: 'I'm a programmer.' Among other things he talks about the history of the famous IBM Big Byte (9 bits), and its final conclusion.

Robert W Bemer, one of the computer industry's 'grand old men', is young and energetic in his appearance and outlook. When a relative newcomer to th industry asked him recently what he did, he replied: 'I'm a behind-the-scenes cat.'

Nine separate and often concurrent 'lives' come to mind for this particular cat. Since 1949 Bob Bemer has been a pioneer, a user, and in his own eyes he is a programmer. In the eyes of his colleagues he is a standards man, a consummate detective and watchdog, the 'father of ASCII', a professional's professional, an eminent author, and perhaps most important, a 'citizen', coming to grips with the problems computers impose on society at a time when many computer people are still trying to evade or ignore them. In addition to these nine lives, he was an IBM man for 6½ years, a Univac man for three, and a GE man from 1965 until GE men became Honeywell men in 1971.

Bob Bemer's computing career began at the Rand Corporation in 1949. Bemer-the-pioneer tells the story of working one morning, after an all-night graveyard shift, on a 604 board to take square root (a task which had never before been done mechanically). 'Then a round little man approached and asked what I was doing. I told him. He then asked about the calculator, and as I answered each question the next one got more difficult and penetrating, until I was really straining every faculty to answer correspondingly.' The round little man turned out to be John von Neumann; Bemer seems to have been asking (and sometimes answering) key computer-related questions ever since.

Bemer-the-user came into being in 1951 at Lockheed Aircraft Company. He organised the computing departments at Marquardt Aircraft and at Lockheed's new Missiles and Space Company. Today, as staff consultant to Honeywell's vice president of Advanced Systems and Technology, he is still a user; one of his primary tasks is to edit the *Honeywell Computer Journal*, a unique publication which is typeset by computer and simultaneously published on microfiche. (The computer in question is Honeywell's software development machine in Phoenix, Arizona. Whenever the journal is delayed, the computer people have to cope with a knowledgeable in-house user who knows which questions to ask. Bemer-the-programmer, who developed several early programming systems, can't be put off with specious excuses.)

Bemer-the-standards-man and Bemer-the-father-of-ASCII are closely related to Bemer-the-citizen. A streak of idealism focuses Bob Bemer's practical experience and inexhaustible energy. It is thus no accident that he is now chairman of the International Standards Organisation's subcommittee on programming languages.

The current issue of his *HIS* journal, for example, carries the saga of ASCII (American Standard Code for Information Interchange), from the birth of Fortran in 1954 and the nine different codes existing in IBM alone in 1959, through the painful process of negotiation of each mark and code, publication of a proposed standard, debate and discussion, further negotiation of certain codes, then publication of *revised* proposed standards, to the false lull of 1963 when people believed that an era of co-operation and goodwill had begun. IBM was expected to support ASCII.

Then came the bombshell — the 8-bit Extended Binary Coded Decimal Interchange Code (EBCDIC, as users know too well) was the internal code in IBM's 360 series, announced in April 1964 (or '64 April', as Bemer would write in his consistent, computer-compatible way, giving an additional boost to Recommendation 2014 of the ISO every time he writes a date). Instead of spending the $5 million more it would have taken (after the $300 million 360 software development) to make the new computers work with ASCII, IBM preferred to let competitors spend a good deal more making their ASCII machines work with the IBM-standard EBCDIC.

By this time Bemer had left IBM for Univac, which has always been one of the leading companies in America and international standards activities. By the time the international standard was accepted by ISO ('on 67 December 22') he had moved on to GE, and the ASCII code is now used (as the ISO character code) in billions of dollars worth of communication and computer gear around the world.

Another life of this energetic behind-the-scenes cat is Bemer-the-detective — still watching over the progress of standards and seeking out potential difficulties. Bemer's boss at Honeywell knew that he was partly responsible for the IBM 360 series (and all its imitators) having an 8-bit byte. (The rest of the credit goes to Fred Brooks, now professor of computer science at the University of North Carolina.) Thus when recent IBM announcements started using the term 'character' instead of IBM's traditional 'byte', Bemer's boss became suspicious, knowing that IBM never does even a trivial thing casually. It was already known that IBM's chairman Frank Cary had told a high US government official that it wouldn't make any difference in IBM's new line if the government used ASCII or EBCDIC, and thus IBM could meet the requirement that ASCII was a federal standard.

'Then it hit me,' says Bemer. 'IBM could use a nine-bit byte on the new line, which was to be announced 75-wise. With this inspiration, I talked it up. Then someone said: "Oh, you mean the Big Byte". He was thereafter less free to talk to me. We checked: sure enough, all of IBM's new production was 9-bit, but they didn't know why.

'Then IBM announced a new magnetic tape unit, featuring a method for correcting any two out of nine bits in each character position.' Bemer was one of the first to react, contending that if *any* two out of nine could be automatically corrected, then 'any' included the parity bit. 'And if the parity bit has lost its function, it isn't needed any more, and you have *nine* useful working bits per character.' At the same time, IBM was objecting to the proposals for channel interface and data communications standards, contending among other things that the 8-bit orientation in each of these was restrictive. Then the company announced its point-of-sale system with a new communications method — one which was not oriented to the 8-bit byte.

'So the whole picture fell in place,' says Bemer-the-detective, 'except that few would believe my guesses.' Rumours reached *Electronic News* and *Datamation* in the States and raised political questions in Europe, where national computer companies would have to spend millions to make similar changes in their architecture. As the commotion over the Big Byte grew louder, IBM's vice president B O Evans wrote a letter to a US government group investigating the ASCII question, and admitted that there would be a need for many and larger character sets for future equipment, particularly for display terminals and photocomposition. 'Just what I had been preaching,' Bemer exults, 'and doing, with our own photocomposition system.'

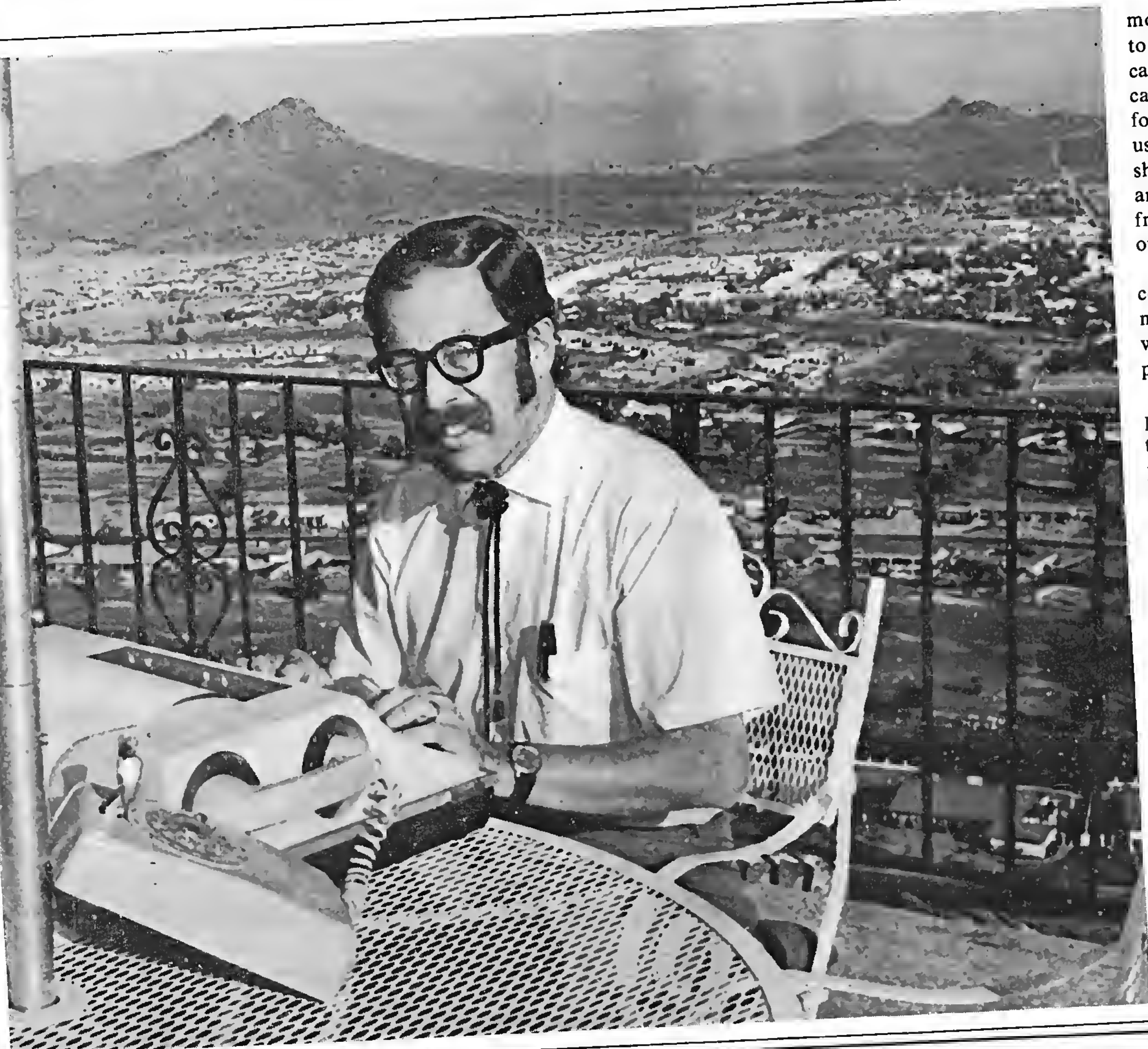
That's the story behind the Big Byte story — one of many in Bemer's career behind the scenes in standards. In this instance if Bob Bemer hadn't fitted the pieces together it might not have been known for another nine months or a year. (IBM is legally constrained, by the 1956 consent decree, *not* to give out information on new products until announcement time.) By then, one wonders what all the 360 and 370 imitators (including not only Siemens but also the entire Eastern Europe Bloc) would have done? IBM would have been able to sail into the new world of using the ISO code, conquer the communications market, and do it all in the name of sweet co-operation, laughing at the imitators and saying gently to users: We found the technology to give you everything, *and* standards, without destroying any little part of your past programs and data files. This may be a good thing for users, but competitors deserve every minute of advance warning they can get. As Bemer says: 'It's diabolic — but so clever that I'm ashamed not to have thought of it myself!'

You don't squeeze this many 'lives' into a standard eight-hour work day. So Bemer often takes home programming tasks. The terminal on his patio is placed to catch the pleasant evening breezes in Phoenix. Working with a high-intensity lamp he gets through a lot of programming and reading. His terminal is linked to any time-sharing service he cares to dial — they all use the ASCII code.

Bemer-the-author has written more than 50 papers, including one landmark paper on time-sharing — in March 1957, while he was working at IBM! 'All that one got me then was a suggestion to my boss that I be fired for promulgating something that wasn't IBM's policy,' he says cheerfully. Another achievement he's proud of is creating the Print 1 programming system for the IBM 705 computer — 'the first load-and-go compiler in history,' says the programmer. 'And the only major software system that IBM ever supplied on pre-announced schedule,' adds Bemer-the-Honeyweller.

Bemer-the-professional might be identified by his honours. He is a Fellow of the British Computer Society (and the only American ever to address a BCS annual general meeting). In the Association for Computing Machinery he began to be active on the editorial board as early as 1957, and was dubbed National Lecturer — a high honour that carries a gruelling schedule with it — in 1961 and again in 1969. He also had the delicate task of organising and planning and chairing the methods and applications programme for the 1973 National Computer Conference in New York last June.

Then in August Bemer appeared on the podium again, at NorData 73 in Copenhagen, discussing the problems of computers and society. Again, instead of glittering generalities, he gave the audience an example-filled description of the harm — and the good — computers can do. 'I like computers,' he said. 'I believe that they are presently



more beneficial than harmful to society, and that this ratio can be increased if we take careful consideration and plan for their best and proper usage. If I were fatalistic, I should feel that they have arrived just in time to save us from our enemies, who are ourselves.

In almost 25 years of being a computer programmer, I have never faced a day of working with computers without pleasant anticipation.

'I also like a fire in the fireplace, but not arson,' he continued. 'Both fire and computers are tools accessible to all of society in some form, and society uses such basic tools in many ways. . . . Any tool that provides leverage or amplification can be misused,' he warned.

Referring to the need for 'computer appreciation and depreciation courses' for citizens, Bemer went on to list the evils arising from the imaginary 'authority' of the computer as a scapegoat'.

On one occasion a Phoenix piano company advertised on the radio that its computer had ordered too much inventory.

Bemer wrote on behalf of ACM offering to fix the program if the piano company would exonerate its computer if it were discovered to be human rather than computer system error.

Thus it was discovered that the piano company didn't even possess or use a computer!

Bemer-the-citizen asks questions when the computer's authority makes it an accomplice to fraud, as in the recent US Equity Funding scandal, or when the machine becomes a sewage system. 'A well-known truism of computer usage is "Garbage In, Garbage Out", but what happens when we put perfectly valid data in,' he asks, 'and can't get it out again without a program to decode it?'

Just as governments have imposed certain safety and amenity requirements upon the manufacture of automobiles, they may impose some requirements concerning computers. 'It seems certain that the computer has a direct effect upon not only the safety of our citizens, but also upon other rights.'

Bemer-the-citizen concludes. 'It might thus be reasonable to demand that software and hardware should also be built to certain specific standards to carefully protect these rights.'

Cores for

Then a round little man approached and asked what I was doing. I told him. He then

... that an era of co-operation and goodwill had begun. IBM was expected to support ASCII.

be announced 7S-wise. With this inspiration, I talked it up. Then someone said: "Oh, you mean the Big Byte". He was

national computer companies would have to spend millions to make similar changes in their architecture. As the commotion over the

... and the technology to give you everything, *and* standards, without destroying any little part of your past programs and data files. This may be a good thing

supplied on pre-announced schedule,' adds Bemer-the-Honeyweller.

Bemer-the-professional might be identified by his honours. He is a Fellow of the

RWB
FYI
& return
&

COMMITTEE CORRESPONDENCE

american national standards committees:
X3—Computers & Information Processing
X4—Office Machines & Supplies

secretariat: CBEMA, 1828 L St., N.W., Washington, D.C. 20036 (202) 466-2299

Doc No. X3L2/1478A
Date 1975 May 19
Project 5
Milestone

RECEIVED
1975 MAY 27
ERIC H. CLAMONS

Reply to: Charles D. Card
Sperry Univac
P. O. Box 500
Blue Bell, Pa. 19422

MEMO TO: X3L2 Members, Alternates, Consultants, and Observers

FROM: Charles D. Card, Chairman X3L2

SUBJECT: Letter ballot on X3L2/1478B
"Proposed Graphics set for the 8-Bit Code"

Attached is X3L2/1478B which represents the consensus reached at Meeting #97 (Marlboro, Mass.) as a basis for the set of graphics to be included in the 8-bit code.

THE BALLOT PERIOD EXPIRES 1975 JULY 7

C. D. Card

We are balloting the following:

"Resolved that the graphics X3L2/1478B be the graphics for use in the 8-bit code."

☐ IN FAVOR

☐ IN FAVOR WITH COMMENTS

☐ AGAINST (explanations of reasons to support this position is expected)

NAME ______________________

SIGNATURE ______________________ DATE __________

☐ COMMENTS OR EXPLANATION ATTACHED

X3L2/1478

NMAH 311/1

GRAPHIC	NAME FOR GRAPHIC
α	Lower Case Greek alpha
β	Lower Case Greek beta
γ	Lower Case Greek gamma
δ	Lower Case Greek delta
ε	Lower Case Greek epsilon
ζ	Lower Case Greek zeta
η	Lower Case Greek eta
θ	Lower Case Greek theta
ι	Lower Case Greek iota
κ	Lower Case Greek kappa
λ	Lower Case Greek lambda
μ	Lower Case Greek mu
ν	Lower Case Greek nu
ξ	Lower Case Greek xi
π	Lower Case Greek pi
ρ	Lower Case Greek rho
σ	Lower Case Greek sigma
τ	Lower Case Greek tau
υ	Lower Case Greek upsilon
φ	Lower Case Greek phi
χ	Lower Case Greek chi
ψ	Lower Case Greek psi
ω	Lower Case Greek omega

GRAPHIC	NAME FOR GRAPHIC
²	Superior 2 (SI units)
³	Superior 3 (SI units)
•	Dot Product
√	Radical/check
♡	Heart Suit Symbol
♠	Spade Suit Symbol
◇	Diamond Suit Symbol
♣	Club Suit Symbol
◆	Radix Point
○	Pi Times, Circular APL Symbol
⌶	I-BEAM APL Symbol
ι	Index Gen, Index of APL Symbol
¬	Logical Not
‘	Single Quote Opening
’	Single Quote Closing

GRAPHIC	NAME FOR GRAPHIC
Π	Upper Case Greek pi
≃	Approximately equal
Σ	Upper Case Greek sigma
∃	At least one exists
Υ	Upper Case Greek upsilon
Φ	Upper Case Greek phi
∂	Partial derivative
Ψ	Upper Case Greek psi
Ω	Upper Case Greek omega
∪	Cup
∩	Cap
∧	AND
∨	OR
‖	Parallel to
└	Lower Left Corner
┌	Upper Left Corner
┐	Upper Right Corner
┘	Lower Right Corner
⊥	Perpendicular to, Decode APL Symbol
‾	Overbar

GRAPHIC	NAME FOR GRAPHIC
+	Cross Junction
¶	Paragraph Mark, Pilcrow
”	Closing Double Quotation Marks
†	Dagger
§	Section Mark
‡	Double Dagger
▭	Rectangle
“	Opening Double Quotation Marks
〈	Opening Angular Bracket
〉	Closing Angular Bracket
±	Plus or Minus
▪	Bullet, Square
—	EM Dash
●	Bullet, Dot
÷	Division
∞	Infinity

GRAPHIC	NAME FOR GRAPHIC
$\equiv$	Identity, Identically equiv.
$\int$	Integral
Γ	Upper Case Greek gamma
Δ	Upper Case Greek delta, Increment
∇	Nabla or del
$\therefore$	Therefore
Θ	Upper Case Greek theta
$\subset$	Is implied by
$\supset$	Implies
Λ	Upper Case Greek lambda
Ξ	Upper case Greek xi
$\times$	Multiplication, Cross Product
$\uparrow$	Tends up to limit, up arrow, take APL symbol
$\rightarrow$	Approaches limit, left arrow, go to APL symbol
$\downarrow$	Tends down to limit, down arrow, drop APL symbol
$\leftarrow$	Right arrow assignment APL symbol
$\geq$	Greater than or Equal to
$\neq$	Not equal to
$\leq$	Less than or Equal to
$^{\circ}$	Degree

Keynote for the second session will be given by Robert W. Bemer, Sr., Honeywell Information Systems, who will speak on "ASCII--The Data Alphabet That Will Endure." ASCII, a 12-year-old data alphabet, has by design the capabilities for expansion and extension not inherent in any other code. It is "best prepared" for the coming fields of networking, electronic funds transfer, text processing and photocomposition, and the automated office.

ASCII - The Data Alphabet That Will Endure

Robert W. Bemer
Honeywell Information Systems, Inc.
Phoenix, Arizona, US

A standard data alphabet is indispensable to understanding communication and reading data in machine-encoded form (not spoken, not written or printed). ASCII (the ISO Code) has, by design, capabilities for expansion and extension not inherent in any other code. The many billions of dollars worth of ASCII-based communication and computation equipment is the best prepared for the coming fields of networking, electronic funds transfer, text processing and photocomposition, the automated office, etc.

The status and prospects of this healthy 12-year-old are explored.

Keywords: Alphabet; ASCII; character; code; ISO; symbol

1. ASCII Is a True Alphabet

Because ASCII [1] and its international identical twin, the ISO Code [2], are actually called "coded character sets" in the formal standards, I must begin by explaining why I use the term "alphabet" instead. One reason is that another international identical twin is called the C.C.I.T.T. Working Alphabet No. 5. For the others, some definitions are necessary. Even though I dislike Webster's Third International (sic) Dictionary intensely, here are some of the things it says that "alphabet" means:

- **1a.** Any particular set of letters with which one or more languages are written, especially such a set of letters arranged in customary order.
- **1b.** Any set of characters with which one or more languages are written, whether these characters are letters (sense 1a), signs of a syllabary, or other basic units of writing.
- **1f.** The alphabetic system of writing, as distinguished from syllabic, ideographic, and other systems.
- **1h.** Any system of signs or signals, visual, auditory, or tactile, that serve as equivalents for the usual written letters of the alphabet.
- **1i.** A particular set of names used to designate the various letters in the alphabet (the pronouncing alphabet used in civil aviation).
- **1j.** In cryptology, a set of one-to-one equivalences between a sequence of plaintext letters and the sequence of their cipher substitutes.

Are not ASCII and the ISO Code actually alphabets in every such sense?

In sense **1a**, it contains letters, and they are arranged in the customary order (but not collating sequence, because of the dual case representations of the letters) by the numerical order of their bit encodings. More than one language can be written with it; as the international code, the most prevalent languages using the Roman alphabet can be written. Note particularly, in the official reference version, that provision is made for the extra Scandinavian letters, located in the proper position (although the usage is not the same in Denmark, Norway, and Sweden).

In sense **1b**, it contains other units of writing. Punctuation is there, as are underscore and other common symbols. Diacritical elements exist for forming compound and accented letters, thus bringing more Roman-based languages within its capability.

In sense **1f**, it certainly does not have syllabic or ideographic characteristics. So it is not excluded from being called an alphabet for these reasons.

In sense **1h**, the encoded representations are the equivalents. In fact, this is the definition of most interest to us. Note that after 90 years of encoding (starting with punch cards for the 1890 Census), Webster's Dictionary fails to give specific status to this manifestation.

Senses **1l** and **1j** may not appear to be pertinent now, but they are there for a reason, and we shall return to them.

2. Why Alphabets?

Alphabets comprise a class of methods to record knowledge for transmittal to others. To transfer knowledge we must transfer information; to transfer information we must transfer data.

There are, of course, other methods using basic elements at higher levels of complexity -- such as syllables, ideographs, etc. However, the primitive ASCII (ISO Code) is *the* worldwide standard for exchange of data and information. It continues to be the standard because the primitives are representable in alternate but related ways.

It is interesting to speculate what might have happened if our forebears had developed the phonograph or tape recorder *before* writing with alphabets. Would we be as deeply into databanks as we are now? Would standards have been developed for speech sounds in analog form, using the computer to discriminate and remove differences in people's voices? Would Confuclus have said that one analog picture was worth a thousand digital words?

One suspects it might be difficult to search such a databank, however. At least it appears that we do not have the methods yet. For example, in reply to the Senate Committee asking for information from the Nixon tapes, the White House says that they have not been classified as to content, and it would take listening to them in their entirety. In other words, a linear realtime search -- and we know how inefficient that can get as databases grow larger.

3. Why ASCII Survives and Grows

My personal history or view of the development of the ISO Code and ASCII [3] tells of millions of dollars of careful international effort and planning spent in its creation. But that is minor, amortized by many billions now invested in communications and computer equipment that operate via ASCII. It also tells of the IBM code called EBCDIC, a result (according to Fred Brooks, one of the chief designers of the 360) of forced announcement before ASCII peripheral equipment could be completed. Although the 360 was *said* to have ASCII capability, it was never realized in the software.

If computers, in substantial portion, do not operate in ASCII as native mode, then why will EBCDIC not be the survivor? Many people, both in and out of government, have blithely assumed that EBCDIC will -- and continue to invest money in software and operations based upon EBCDIC.[4] They are going to be very surprised, because *IBM knows* that EBCDIC will eventually be subordinatel [5]

Why? Because EBCDIC is not, like ASCII, the result of meticulous design.

> "... it would appear that no single 'computer code' can be completely adequate, and that insistence on a single code for all purposes would be counterproductive. Rather, the Federal Government should maximize the benefits to be accrued from taking advantage of our growing technological ability to live in a multi-code world ..."[5]

In ASCII, the controls are all located in the leftmost two columns. It is compact, extensible, expandable, and even subsettable. It can grow easily into an 8-bit code (expandability), or into 9-bits, 10-, or anything. At any level of byte size, it can be extended to encompass alternate sets of characters; keeping the same control columns, various pages can be substituted for the other columns. The methods for expansion and extension are also standardized.[6,7] Sets having sufficient utility may be registered for international usage, via the French standards body AFNOR, which holds the secretariat for international code standardization within ISO TC97. The vehicle for doing this is the ESCape character.[8] Various pages are registered with unique ESCape sequences.[9,10]

4. Code Extension

In the extension procedures, the existing 7-bit ASCII is divided into control and graphic portions. The first two columns of code -- the controls -- comprise the C0 set; the other six columns --the graphics -- comprise the G0 set. The extended set first removed from basic ASCII is similarly divided into the C1 and G1 sets: Obviously such sets could be adjoined in the 8-bit form, and the USSR [11] and Japanese [12] standards are excellent examples of so doing.

4.1 Extended Control Sets

C1 sets can be, and have been, designed for many purposes. The one furthest progressed to agreement is that for softcopy controls, for CRT display screens [13,14].

> "... The major difficulties at present are in using the established 'control' characters with devices that had not been invented at the time the code was, or in extending the 'graphic' symbol set to meet new application requirements ..."[5]

Work is continuing in both ECMA (the European Computer Manufacturers Association) and ANSI X3 to get agreement sufficient for final registration. The original work of X3L2 was for softcopy controls to be in an *expanded* set (8-bit code), but that is presumptious.

b_4	b_3	b_2	b_1	b_7 b_6 b_5 →	000	001	010	011	100	101	110	111
					0	1	2	3	4	5	6	7
0	0	0	0	0			SP	0		P		
0	0	0	1	1				1	A	Q		
0	0	1	0	2			"	2	B	R		
0	0	1	1	3				3	C	S		
0	1	0	0	4			$	4	D	T		
0	1	0	1	5				5	E	U		
0	1	1	0	6				6	F	V		
0	1	1	1	7				7	G	W		
1	0	0	0	8			(	8	H	X		
1	0	0	1	9			)	9	I	Y		
1	0	1	0	10			*		J	Z		
1	0	1	1	11			+	;	K			
1	1	0	0	12			,	<	L			
1	1	0	1	13			−	=	M			
1	1	1	0	14			.	>	N			
1	1	1	1	15			/		O			

Figure 1. COBOL Character Set

b_4	b_3	b_2	b_1	b_7 b_6 b_5 →	000	001	010	011	100	101	110	111
					0	1	2	3	4	5	6	7
0	0	0	0	0			SP	0		P		
0	0	0	1	1				1	A	Q		
0	0	1	0	2				2	B	R		
0	0	1	1	3				3	C	S		
0	1	0	0	4			$	4	D	T		
0	1	0	1	5				5	E	U		
0	1	1	0	6				6	F	V		
0	1	1	1	7			'	7	G	W		
1	0	0	0	8			(	8	H	X		
1	0	0	1	9			)	9	I	Y		
1	0	1	0	10			*	:	J	Z		
1	0	1	1	11			+		K			
1	1	0	0	12			,		L			
1	1	0	1	13			−	=	M			
1	1	1	0	14			.		N			
1	1	1	1	15			/		O			

Figure 2. Fortran Character Set

ISO Technical Committee 46 (Documentation), in its Subcommittee 4 (Automated Documentation) has a working group on bibliographic codes. Its first candidate for registration as a C1 control set is a set for bibliographic controls [15] to be embedded in text to delimit certain special data. This C1 set contains four classes of characters -- annotation controls, filing controls, reference controls, and subject designators. Major credit should be given here to Dr. Ernst Kohl of the Bavarian State Library in Munich.

Although little work has been done, other C1 control sets are envisioned for typographic control -- to vary the font, weight, slope, size, and spacing, etc. of the graphic characters. Other sets could be envisioned for fields such as process control, animation and other graphics applications, sewing machines, etc. Do you think the last one far-fetched? Singer has already announced a machine with a microcomputer, and is there any reason to think that future models won't use ASCII characters, in a hand calculator type of display, to give instructions and options available?

4.2 Extended Graphic Sets

G1 sets are further along. ISO TC46/4/1 has tabled Draft International Standards for Latin (DIS-5426), Greek (DIS-5427), and Cyrillic (DIS-5428). We may presume the latter is in harmony with [11]. Under study are sets for mathematical characters and the African languages. Proposals have been solicited for such languages as Arabic, Kata Kana, Kanji, etc. It is permissible for a G1 set to be a partial replication of the basic G0 set of ASCII; indeed, many are very similar, with the lower case being replaced by the new alphabet.

Although we have seen the ISO assignments for natural languages to be in the jurisdiction of TC46, TC97 (Computers and Information Processing) has retained authority to make assignments for programming languages. The work has been concentrated in G1 sets for COBOL (figure 1), Fortran (figure 2), Basic (figure 3), and PL/I (figure 4). A table for APL is being constructed. ALGOL presents different problems.[16]

b	b	b	b		0	1	2	3	4	5	6	7
				b	0	0	0	0	1	1	1	1
				b	0	0	1	1	0	0	1	1
				b	0	1	0	1	0	1	0	1
0	0	0	0	0			SP	0		P		
0	0	0	1	1			!	1	A	Q		
0	0	1	0	2			"	2	B	R		
0	0	1	1	3			#	3	C	S		
0	1	0	0	4			$	4	D	T		
0	1	0	1	5			%	5	E	U		
0	1	1	0	6			&	6	F	V		
0	1	1	1	7			'	7	G	W		
1	0	0	0	8			(	8	H	X		
1	0	0	1	9			)	9	I	Y		
1	0	1	0	10			*	:	J	Z		
1	0	1	1	11			+	;	K			
1	1	0	0	12			,	<	L			
1	1	0	1	13			-	=	M			
1	1	1	0	14			.	>	N	^		
1	1	1	1	15			/	?	O	_		

Figure 3. BASIC Character Set

b	b	b	b		0	1	2	3	4	5	6	7
				b	0	0	0	0	1	1	1	1
				b	0	0	1	1	0	0	1	1
				b	0	1	0	1	0	1	0	1
0	0	0	0	0			ƀ	0		P		
0	0	0	1	1			|	1	A	Q		
0	0	1	0	2				2	B	R		
0	0	1	1	3				3	C	S		
0	1	0	0	4			$	4	D	T		
0	1	0	1	5				5	E	U		
0	1	1	0	6			&	6	F	V		
0	1	1	1	7			'	7	G	W		
1	0	0	0	8			(	8	H	X		
1	0	0	1	9			)	9	I	Y		
1	0	1	0	10			*	:	J	Z		
1	0	1	1	11			+	;	K			
1	1	0	0	12			,	<	L			
1	1	0	1	13			-	=	M			
1	1	1	0	14			.	>	N	¬		
1	1	1	1	15			/		O	_		

Figure 4. PL/I Character Set

One may be tempted to think of these not as G1 sets but rather as subsets of the G0 set, standard ASCII. But note that they are incompatible in minor ways, particularly for PL/I. which was the cause of considerable difficulty in stabilizing ASCII. So perhaps the G1 status is an easy solution.

4.3 The Registry Method

A responsible standardizing body with a specific proposal makes application to AFNOR, acting as agent for ISO TC97/SC2. Applications may be for graphic sets (G1, etc.), C0 or C1 control sets, a single control character, or a code requiring special interpretation. The approval procedure is defined in [10]. A unique ESCape sequence is assigned. It is here that definitions *1/* and *1/j* for alphabet become applicable. The ESCape sequence, as adjoined to any following character before termination, becomes a name for the alternate characters and alphabet, in one-to-one equivalence. Thus all of the world's symbol and alphabets may be represented uniquely for interchange.

5. How IBM Can and Will Use ASCII

EBCDIC is a sparsely settled code that utilizes the 8-bit capability of 256 characters ineffectively. The collating sequence(s) are not easily derivable from the numerical values of the coded representations. The controls are intermingled with the other characters, so that it can not be extended by paging, as ASCII can. It has only one redeeming virtue -- one-to-one correspondence with ASCII via a common character set as represented in punch cards! (See figures 5, 6)

> "The interesting observation is that if two character codes each have the same symbol set, and if each meet the requirement of no symbol ambiguities for the same bit pattern (no duals), then automatic context-free translation between the two character codes is a trivial task ... The operating cost of translation (between two such character codes) concurrently with preparing or accepting an interchange message is trivial in today's systems and will be more so in tomorrow's LSI machines."[5]

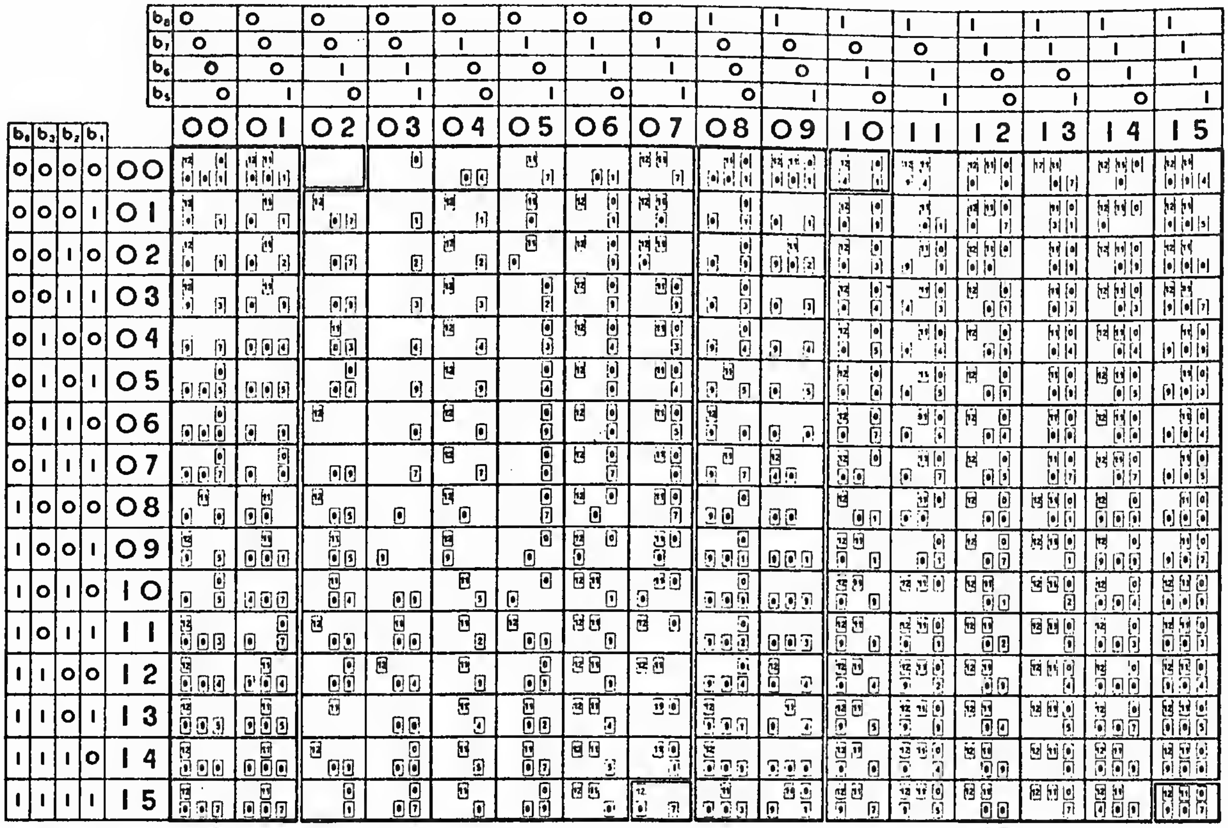


Figure 5. Hole Patterns Assigned By Code

	12	11	0		12 0	12 11	11 0	12 11 0	12	11	0		12 0	12 11	11 0	12 11 0	
	02/06	02/13	03/00	02/00	07/11	07/12	07/13	11/10	10/08	11/01	11/09	06/00	12/03	12/10	13/01	13/08	8 1
1	04/01	04/10	02/15	03/01	06/01	06/10	07/14	13/09	00/01	01/01	08/01	09/01	10/00	10/09	09/15	11/11	9 1
2	04/02	04/11	05/03	03/02	06/02	06/11	07/03	13/10	00/02	01/02	08/02	01/06	10/01	10/10	11/02	11/12	9 2
3	04/03	04/12	05/04	03/03	06/03	08/12	07/04	13/11	00/03	01/03	08/03	09/03	10/02	10/11	11/03	11/13	9 3
4	04/04	04/13	05/05	03/04	06/04	06/13	07/05	13/12	09/12	09/13	08/04	09/04	10/03	10/12	11/04	11/14	9 4
5	04/05	04/14	05/08	03/05	06/05	08/14	07/06	13/13	00/09	08/05	00/10	09/05	10/04	10/13	11/05	11/15	9 5
6	04/06	04/15	05/07	03/06	06/06	06/15	07/07	13/14	08/06	00/08	01/07	09/06	10/05	10/14	11/06	12/00	9 6
7	04/07	05/00	05/08	03/07	06/07	07/00	07/08	13/15	07/15	08/07	01/11	00/04	10/06	10/15	11/07	12/01	9 7
8	04/08	05/01	05/09	03/08	06/08	07/01	07/09	14/00	09/07	01/08	08/08	09/08	10/07	11/00	11/08	12/02	9 8
9	04/09	05/02	05/10	03/09	06/09	07/02	07/10	14/01	08/13	01/09	08/09	09/09	00/00	01/00	08/00	09/00	9 8 1
8 2	05/11	05/13	05/12	03/10	12/04	12/11	13/02	14/02	08/14	09/02	08/10	09/10	14/08	14/14	15/04	15/10	9 8 2
8 3	02/14	02/04	02/12	02/03	12/05	12/12	13/03	14/03	00/11	08/15	08/11	09/11	14/09	14/15	15/05	15/11	9 8 3
8 4	03/12	02/10	02/05	04/00	12/06	12/13	13/04	14/04	00/12	01/12	08/12	01/04	14/10	15/00	15/06	15/12	9 8 4
8 5	02/08	02/09	05/15	02/07	12/07	12/14	13/05	14/05	00/13	01/13	00/05	01/05	14/11	15/01	15/07	15/13	9 8 5
8 6	02/11	03/11	03/14	03/13	12/08	12/15	13/06	14/06	00/14	01/14	00/06	09/14	14/12	15/02	15/08	15/14	9 8 6
8 7	02/01	05/14	03/15	02/02	12/09	13/00	13/07	14/07	00/15	01/15	00/07	01/10	14/13	15/03	15/09	15/15	9 8 7

Figure 6. Code Assigned To Hole Patterns

Anyone with a $15 hand calculator realizes how cheap a microelectronic chip must be, for his calculator does a more complex job than the job of converting back and forth between the ASCII and EBCDIC encodings for the same character.

So let us postulate a very inexpensive chip inside IBM computers. It converts from EBCDIC to ASCII or from ASCII to EBCDIC without any delay as input-output or other operations are executed. Two questions must be asked:

- Is the data EBCDIC or ASCII?
- Does the program expect EBCDIC data or ASCII data?

Imagine subscripting the comparison instruction by EBCDIC or ASCII tags. I have a master file in EBCDIC, against which I run an update tape in ASCII. My program says "Compare the keys on an ASCII basis". The CPU, noting that the key of the master file is in EBCDIC, routes it through the chip before attempting to compare it to the key from the update record. My program can also give instruction to convert the entire updated file to ASCII automatically as it is being stored.

Thus the feasible technique. How about some signs that it will be so?

- Note IBM's commitment to word processing and photocomposition. According to B. O. Evans:

"There are similar requirements for the ability to use differing codes in differing contexts to represent different graphic needs. In particular, the development of 'end user' devices, such as photocomposers and interactive displays, requires greatly expanded symbol sets to be developed for some applications -- more symbols than can be contained directly in a 7- or 8-bit code. Thus, specific codes (of however many bits are required to represent a symbol) might well be developed for data interchange in certain application areas utilizing such devices."[5]

- Note IBM's 6250 cpi magnetic tape, which departs from cross-tape parity checking and recognizability of code without programmed knowledge.

- Note IBM's firm and continuing insistence that data control procedures be bit-transparent and not byte-oriented -- particularly not 8-bit-byte-oriented.

"We expect to see machine architectures having the flexibility to adapt efficiently to as-yet-undefined code structures without disruption of existing applications at any time even after the system was installed."[5]

- Note SDLC and SNA. When you are going to be communications-oriented, and even run a satellite system, why object to using the ASCII code that *all* communications is based on?

6. Conclusions

ASCII was well-designed, and is flexible to adapt to usage that may take any turn of development. There is no reason, with presently available technology, to use any other encoded alphabet. As a single standard, it enables private data to become public whenever that is desirable (i.e., privacy may be protected or maintained in ways other than unintelligibility). It is the alphabet of all communications networks, of all minicomputers, and of some larger computers such as the NCR Century series.

After twelve years it is still healthy, and when IBM puts the seal of approval on it (viz. virtual memory and APL) it will be the undisputed universal interchange medium and linguist.

7. References

1. ANSI X3.4 - 1974, "American national standard code for information interchange"
2. ISO 646 - 1973, "6 and 7 bit coded character sets for information processing interchange"
3. R.W.Bemer, "A view of the history of the ISO code", Honeywell Comput. J., 6, No. 4, 274-286 (1972)
4. Federal Information Processing Standard Publication 7, "Implementation of the code for information interchange and related media standards"
5. B.O.Evans, letter of 1973 June 28 to R.R.Johnson, Chairman Federal Information Processing Standards Task Group 12.
6. ANSI X3.41 - 1974, "Code extension techniques for use with the 7-bit coded character set of (ANS) code for information interchange"
7. ISO 2022 - 1973, Code extension procedures for ISO 7 bit code"
8. R.W.Bemer, "Escape - a proposal for character code compatibility", Commun. ACM 3, No. 2, 71-72 (1960 Feb)
9. R.W.Bemer, Working Paper, "Registry - A new feature for ISO recommendations", 1970 May 28
10. ISO 2375 - 1974, "Procedures for the registration of escape sequences"
11. USSR State Standard, GOST 13052-67, "Computers and data transmission equipment, alpha-numeric codes", 1967 July 10
12. Japanese Industrial Standard Code for Information Interchange, "JISCII", C6220-1969
13. T.O.Holtey and E.H.Clamons, "Soft-copy controls", Honeywell Comput. J. 7, No. 4, 267-269 (1973)
14. ANSI X3L2/1499, "Proposed American national standard - control characters for an 8-bit code", 1975 Oct 10
15. ISO/TC46/4/1 N54, "The set of specific bibliographic control characters", 1975-08-12
16. ISO Third Draft Proposal 1672, "Hardware representation of ALGOL basic symbols in the ISO 6 and 7 bit coded character sets"