R. W. Berner

Speeches and Papers

1977-

1978

1978

Computer History Museum

SEPARATION SHEET
Computer History Museum

Category of item(s):

- ☐ Audiocassette
- ☐ Electronic records
- ☐ Oversized materials
- ☒ Photographs/negatives
- ☐ Printed materials
- ☐ Three-dimensional objects
- ☐ Videotape
- ☐ Other: _____

Number of items: 3 sheets of slides (35mm) - 55 slides

Provide information about item(s):

Presentation slides. There is no clear title but is related to international standards. Found in the 1978 part of the folder, so likely dated 1978.

Originally filed in:

Catalog #: 102785430 Lot #: X3054.2005

Collection Name: Robert (Bob) Bemer papers

Series: Papers Subseries: _____

Folder title: Speeches and papers, 1955-1982
Folder 4 of 5

Relocated to:

Catalog #: 102785615 Lot #: X3054.2005

Collection Name: Robert (Bob) Bemer papers

Series: Papers Subseries: _____

Folder title: Presentation slides, 1970-1978

Purpose of removal: To separate the slides from the text.

Date: 2/28/19

Name: Sydney Olson

1434t

RWB ERRATA COPY
RED = MANDATORY CHANGES
BLUE = DESIRABLE CHANGES

INSIDE ASCII

By R. W. Bemer

The data alphabet called ASCII (Figure 1, page 98, and Reference 1), also has two other names—International Standard 646 (the ISO Code [Reference 2]) and Alphabet No. 5 of CCITT (the International Consultative Committee for Telephone and Telegraph). It is used throughout the world, incorporated in billions of dollars of equipment.

But is it used correctly and wisely? Not always. There are misinterpretations, and gaps in definition that permit nonstandard usage. This article (in three parts) will give you the background, peculiarities, preferred practices, and new developments for ASCII. You will find a lot of information not too generally known or realized; it should help in the correct and safe usage of ASCII. For additional help, you can reference the various national and international standards given in Table 1. Some other detailed articles are listed in References 3, 4 and 5.

	ISO	ECMA	ANSI	FIPS PUB	CSA	BS	AS	CCITT	JIS	GOST
Binary-coded Character Set	646	6	X3.4-1977 $4.50	1	Z243.4	4730	1776	V.3	C6220	13052-67
Graphics for Control Characters	2047	17	X3.32-1973 $3.50	36		4730				
Character Set for Handprinting	97/3 WG9		X3.45-1974 $3.75	33	Z243.34.1					
Additional Controls Character Imaging		48	DIS $3.64							
6-bit Sets	963	14		15	Z243.6	4731/1	1070			
Code Extension Techniques	2022	35	X3.41-1974 $6.00	35	Z243.35	4953				
Registration Procedures for Escape Sequences	2375									
8-bit Coded Character Set	DIS 4873	43	X3.2/77/08							
Character Set for 7 x 9 Matrix Printers		42								
Keyboard	2530	23	X4.14-1971 $3.75					4822/1	1922	
Character Sets for Programming Languages	97/5 N436	53								

Legend

ISO - International Standards Organization
ECMA - European Computer Manufacturers Association
ANSI - American National Standards Institute
FIPS - Federal Information Processing Standard
CSA - Canadian Standards Association
BS - British Standard
AS - Australian Standard
CCITT - Consultative Committee International, Telephone & Telegraph
JIS - Japanese Industrial Standard
GOST - USSR Standard

Table 1.

STICKS 4-7

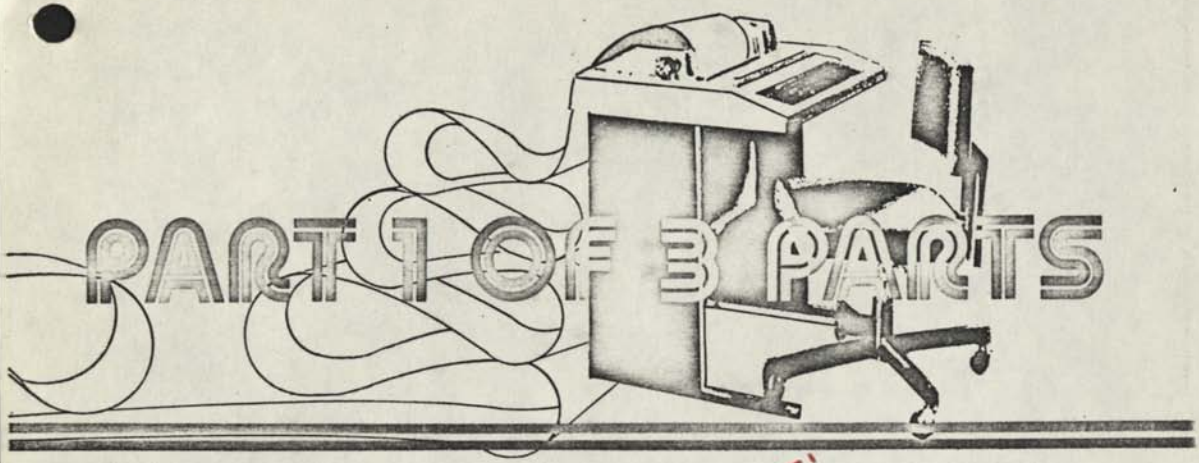
ASCII, as a 7-bit code, is usually represented in 8 columns of 16 positions. The row positions are 0000 through 1111, the low-order 4 bits, 0 through 15 in decimal. The columns are 000 through 111, the next higher 3 bits, 0 through 7 in decimal. For some reason, the developers of ASCII found it convenient to refer to these eight columns as "sticks." So shall we. Each position will be represented in this article by its usual decimal representation. For example, capital A is position 4/1. Figure 2 is a representation of ASCII that is more convenient to those working in octal, rather than hexadecimal, notation.

HIGH ORDER OCTAL DIGITS →	00	02	04	06	10	12	14	16	← LOW ORDER OCTAL DIGIT
	NUL	DLE	SP	0	@	P	`	p	0
	SOH	DC1	!	1	A	Q	a	q	1
	STX	DC2	"	2	B	R	b	r	2
	ETX	DC3	#	3	C	S	c	s	3
	EOT	DC4	$	4	D	T	d	t	4
	ENQ	NAK	%	5	E	U	e	u	5
	ACK	SYN	&	6	F	V	f	v	6
	BEL	ETB	'	7	G	W	g	w	7

HIGH ORDER OCTAL DIGITS →	01	03	05	07	11	13	15	17	← LOW ORDER OCTAL DIGIT
	BS	CAN	(	8	H	X	h	x	0
	HT	EM	)	9	I	Y	i	y	1
	LF	SUB	*	:	J	Z	j	z	2
	VT	ESC	+	;	K	[	k	{	3
	FF	FS	,	<	L	\	l	|	4
	CR	GS	-	=	M	]	m	}	5
	SO	RS	.	>	N	^	n	~	6
	SI	US	/	?	O	_	o	DEL	7

Figure 2. ASCII IN OCTAL REFERENCE FORM

TABLE 1a. LOGICAL SUBSETS FOR ASCII



The first positions of sticks 4 and 6 are respectively the "commercial at" and "accent grave." Then the upper and lower case Roman alphabets follow. This offset of one position is historical (from the United Kingdom), and of no importance as long as you remember that it is so.

Following the alphabet in both sticks 5 and 7 are three positions each that one must be very cautious about. In ASCII they are assigned as [, /¹, and] in stick 5 — {, |², and } in stick 7. But in the ISO Code and CCITT versions they are reserved for national usage. Table 2 gives the national use assignment for these positions. Surely you remember that the Scandinavian alphabet has 29 letters, not 26? My friend Örjar Heen in Oslo is very protective of these positions. He says "If you Americans want to sell computers and software abroad, don't use the ASCII characters for these positions in your software."

To be more precise, positions 5/11, 5/12, 5/13, 7/11, 7/12, and 7/13 (noted above) are called *primary* national usage positions. So is 4/0, where ASCII has the "commercial at." Honeywell, for example, uses the "at" in its timesharing systems for deleting the previous character upon entry. But this isn't too serious, because many nations also have the "at" in their primary sets.

	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	#	$	@	IJ	½	'	^	`	é	ù	è	¨
Australia	#	$	@	[	\	]	^	`	{	|	}	~
Belgium–A	£	$	@	°	Ç	§	^	`	é	ù	è	¨
W. Germany–A	#	$	@	Ä	Ö	Ü	^	`	ä	ö	ü	ß
US	#	$	@	[	\	]	^	`	{	|	}	~
Japan	#	$	@	[	¥	]	^	`	{	|	}	~
UK	£	$	@	[	\	]	^	`	{	|	}	~
Italy–A	£	$	@	°	ç	§	^	`	é	ù	à	ò
Switzerland–A	ù	ä	à	ö	ü	è	^	`	é	ï	ä	ö
France–A	#	$	à	°	ç	§	^	`	é	ù	è	¨
USSR	#	;	Ю	Б	Ц	Э	^	`	Ъ	Ы	Ь	Ё
Netherlands–B	#	$	@	IJ	½	'	^	`	é	ù	è	¨
Belgium–B	£	$	@	°	Ç	§	^	`	é	ù	è	¨
France–B	£	$	à	°	ç	§	^	`	é	ù	è	¨
Switzerland–B	ù	ä	à	ö	ü	è	^	`	é	ï	ä	ö
Italy–B	£	$	@	°	ç	§	^	`	é	ù	à	ò
Switzerland–C	ù	ä	à	ö	ü	è	^	`	é	ï	ä	ö
Hungary	#	Ü	É	Á	Ö	Ü	Ó	Ö	é	ú	ü	ó
W. Germany–B	#	$	@	Ä	Ö	Ü	^	`	ä	ö	ü	ß
Switzerland–D	è	ä	à	é	ö	ü	^	`	é	ï	ä	ö
Sweden	#	¤	É	Ä	Å	Ö	Ü	é	ä	å	ö	ü
Finland	#	¤	@	Ä	Å	Ö	Ü	é	ä	å	ö	ü
Denmark	#	$	@	Æ	Ø	Å	^	`	æ	ø	å	~
Norway	#	¤	@	Æ	Ø	Å	^	`	æ	ø	å	~
Spain	#	$	@	¡	Ñ	¿	^	`	{	ñ	}	~

Table 2. National Usage

Also in sticks 4-7 are three diacritical marks. They are accent grave (`) in 6/0, circumflex (^) in 5/14, and tilde (~) in 7/14. These are called *secondary* national usage positions. In some countries the tilde is a straight overline.

But it is the circumflex where we have a lot of confusion. Teletype first made it an "up arrow" in an earlier version of ASCII, to serve as an exponentiation symbol, primarily for BASIC. But that doesn't do very well, because the exponentiation for FORTRAN is a double asterisk! The FORTRAN version is preferable in France, certainly, because they use such words as crane, côte, cout, and so on.

A companion problem exists in position 5/15, with the underscore. The underscore is neither national nor diacritical; all countries use it just as underscore (and for typesetting it is a U.S. convention to indicate italics, but in Italy it means boldface, except when it is the last character in a line!³). But Teletype's early version of ASCII used it as a "left arrow" — probably for an assignment symbol equivalent to := in ALGOL. The up and left arrow have been carried over from Teletype into many video terminals. Ask your terminal manufacturer to cease and desist and retrofit. It's not ASCII and will only cause trouble forever.

The last character in sticks 4-7 is the Delete, symbol DEL, in position 7/17. It was put here because the binary code is 1111111, which would be all punched holes in perforated (not always paper!) tape, and that is the only way to make sure that it cannot be misread as some other character. ASCII is a complete set; all positions are assigned to have meaning.

STICKS 2-3

These are usually called the sticks for digits and specials. Remember that they are the "digits" 0 to 9; not numbers, not numerals, not anything but digits! They are in 3/0 through 3/9 so that the low-order 4 bits are the representations for packed decimal. Originally we considered the possibility of a special 4-bit set for numerical applications (see the fifth entry in Table 1a), but it turned out that computer hardware became inexpensive enough to not deprive ourselves of the extra capabilities of the 7-bit and 8-bit sets.

b_4 b_3 b_2 b_1	ROW	COL 0000 0	COL 0001 1	COL 0010 2	COL 0011 3	COL 0100 4	COL 0101 5	COL 0110 6	COL 0111 7
0000	0	NUL	DLE	SP	0	NOTE 1	P	NOTE 1	p
0001	1	SOH	DC1	!	1	A	Q	a	q
0010	2	STX	DC2	"	2	B	R	b	r
0011	3	ETX	DC3	NOTE 1 £ #	3	C	S	c	s
0100	4	EOT	DC4	NOTE 1 $	4	D	T	d	t
0101	5	ENQ	NAK	%	5	E	U	e	u
0110	6	ACK	SYN	&	6	F	V	f	v
0111	7	BEL	ETB	'	7	G	W	g	w
1000	8	BS	CAN	(	8	H	X	h	x
1001	9	HT	EM	)	9	I	Y	i	y
1010	10	LF	SUB	*	:	J	Z	j	z
1011	11	VT	ESC	+	;	K	NOTE 1 [	k	NOTE 1
1100	12	FF	FS	,	<	L	NOTE 1 \	l	NOTE 1
1101	13	CR	GS	−	=	M	NOTE 1]	m	NOTE 1
1110	14	SO	RS	.	>	N	NOTE 1 ^	n	NOTE 1
1111	15	SI	US	/	?	O	—	o	DEL

Note 1

These 12 positions are variable for national usage — 2 for currency, 7 primary national usage, and 3 secondary usage which are diacritical marks when preceded by BSP. The presently known assignments are given in the table below.

Figure 1. ASCII (ISO Code)

Position 2/0 is officially called "space." I don't and didn't like it, and would have preferred "blank." Which is why the IBM community often uses a lower case "bee" with a slash through the vertical as its symbol. From the Univac side, the space has the official symbol "delta."

Having mentioned packed decimal, where two digits go into each 8-bit group ("byte" to the American, "octet" to the French), a word of caution on the plus and minus signs — they are in stick 2, rather than stick 3 with the digits. But the low order 4 bits are distinct, and + should be used only as 1011, – only as 1101. I mention this because the nonstandard code EBCDIC permits multiple representations of + and – in packed decimal. And the ASCII representations are not even coincident with any of these, with obvious dangers!

Watch out for the "currency" positions, 2/3 and 2/4. They also have national variations. In ASCII they are customarily # and $, but there are some things to be remembered:

- # is not "number sign" for many countries, most of which use "No." or "Nr." for that purpose. And when it is "number," it must precede the digits, not follow.
- # closely resembled the "sharp sign" in music.
- # is "pound sign" only for the U.S., the only major country still not using the metric system. To the rest, it's kilograms. For now, it's best to use the abbreviation "lb." in the U.S., not the #. In any case, both must *follow* the numeral.
- To the British, a "pound" has the symbol "£", which is why that is the symbol in position 2/3 for the UK. They get very irked when # is called a "pound" sign, especially in software manuals.
- The "dollar" is peculiar to the U.S., Canada, and some others. There are also francs, marks, escudos, pesos, lire, etc., etc. Which is why the ISO code uses the universal currency symbol in position 2/4. It's a circle with outside spikes at 45, 135, 225, and 315 degrees (¤), called "scarab." Table II also shows these assignments for several countries.
- ECMA has provided a separate guideline for specifying international currencies. See the "Where to Get More Information" at the end of this article.

It's a tough problem, and will get worse when we get into expanded character sets for photocomposition and such. For now, all we can do is follow the ASCII standard, which says that # is a "number sign."

Only a few more peculiarities remain for sticks 2-3. An important one is in the double quote, position 2/2, and the single quote, position 2/7. That is, you may think it is a single quote, and even use it so, but it is really an "accent acute" for vowels. It slants from top right to bottom left, to complement "accent grave" in 6/0, which slants from top left to bottom right. Some terminal makers do not realize this pairing, and will have accent grave slanting correctly, but put accent acute as a single quote in the unstylized up and down method. My Terminet is one of those that is OK.

Don't forget that to the typesetter, in contrast to typewriters, both single and double quotes have two forms — opening and closing. In fact, the typesetter gets his double quotes by using two single quotes, of either form, because the quote uses very little space in variable space typesetting. Most terminals, either video or hardcopy, use constant spacing. So double and single quotes must be distinct for that reason.

The last variation is in position 2/6, the ampersand. There are many legitimate different symbols for the ampersand. Neither ASCII nor the ISO Code prescribe any particular one. But this leads us to the next topic — how to represent the ASCII characters in handprinted form, so that they may be input to computer systems.

HANDPRINTING FOR STICKS 2-7

The classical confusion for many years was between the digit zero and the letter "oh," but there are other possibilities for confusion. American Standard X3.45 specifies the handwritten character shapes shown in Figure 3.

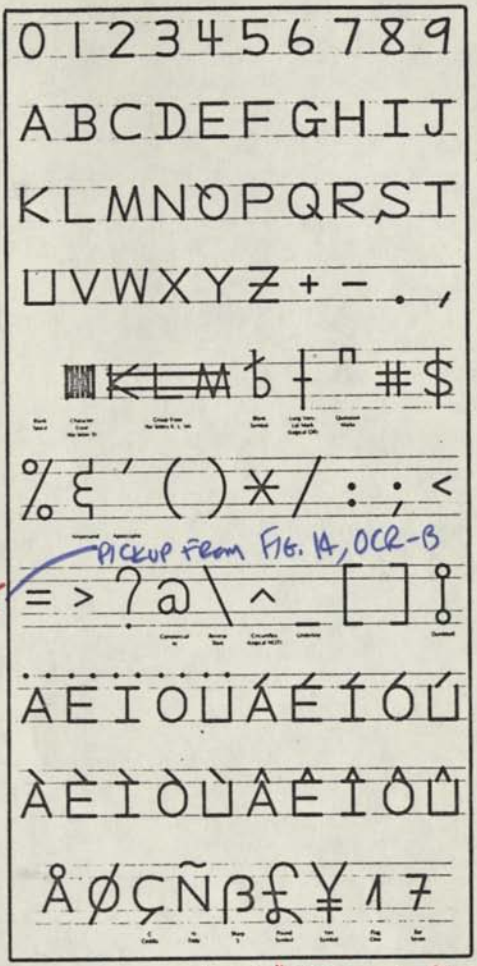


Figure 3. HANDPRINTING FOR ASCII CHARACTERS

This clears up a longstanding problem. The communications types, and the armed services, used to put a slash through the zero; somehow the IBM users got to putting the slash through the letter "oh" instead, confusing the Scandinavians greatly. Now it's neither (which helps), just a 180-degree rotation of the letter Q. The earlier German Standard DIN 66 002 prescribed the cursive loop in the upper right, as some may have learned in penmanship courses. It now permits the ANSI form as well.

UPPER AND LOWER CASE LETTERS

Many people are accustomed to using upper case only. This is a hangover from early line printers and limited sets (until the Stretch computer of IBM, characters were usually 6 bits in size). It would have been far better if they had all been lower case in those smaller sets. Putting it simply, would you buy a book to read if it were all

in upper case? Because lower case is much easier and faster to read, lower case should be the default case when one has only the one case. There is no reason why FORTRAN or BASIC processors cannot understand lower case variable names and verbs just as easily as they can understand upper case.

I always recommend getting a terminal with both cases if it is at all affordable. Second best is making sure that a single-case terminal is retrofittable later, if necessary. And if a single-case terminal, get it in lower case only, if possible. There has been much reportage in the computer trade press about eyestrain resulting from using computer terminals. Is the reason obvious?

STICKS 0, 1

These are the control characters. The most important distinction in ASCII is the split between sticks 0-1, Controls, and sticks 2-7, Graphics. We'll see this later on in the standards for Code Expansion (to 8 bits or more), and Code Extension (alternate sets, such as Cyrillic for the USSR, and Kata Kana for Japan).

Unfortunately, there is, despite the standard, much difference between the ways that various terminal devices handle these control characters. They may act differently, or they may not be operative at all. I have two very useful programs, written in the TEX language (Reference 6). One lists each symbol by name and then shows its action between parentheses. The other asks you to depress in turn all the funny keys on your terminal, and then tells you what control character(s) they generate, if any.

GRAPHICS FOR THE CONTROLS

There are standard graphical representations for the 32 controls, space, and delete. They are defined by ISO 2047, American Standard X3.32, and ECMA-17, and are shown integral to Figure 1. Some terminals are advertised as ASCII terminals, and yet generate Greek or other characters for these positions. Don't believe it! These symbols are every bit as useful as any Greek characters could be.

There are five groups in the basic control set.

STICKS 0, 1 — Logical Communication Control (10)

This group is used for both communication and for labeling of media. It includes:

SOH (0/1) (Start of Heading) — used as the first character in the heading of an information message.

STX (0/2) (Start of Text) — terminates the heading just before the text.

ETX (0/3) (End of Text) — Last character in the text message. Unfortunately, it is generated on many terminals via Control-C, and that's just to the right of Control-X on the keyboard, which is commonly used to cancel a bad input line. And if you mis-key — ouch!

EOT (0/4) (End of Transmission) — the last character in any transmission, and usually it turns your device off!

ENQ (0/5) (Enquiry) — requests a response from a remote station, either an identification of that stations (Who are you?) or its status.

ACK (0/6) (Acknowledge) — used by a receiver to reply "yes" to a sender.

DLE (1/0) (Data Link Escape) — an Escape character, especially for communications, analogous to ESC (1/11). It signals the start of a character sequence that causes a shifting into another set of communication controls, whenever they are needed.

NAK (1/5) (Negative Acknowledge) — used by a receiver to reply "no" to a sender.

SYN (1/6) (Synchronous Idle) — needed by synchronous transmission systems to get into, or stay in, synchronization when no other such signal is available to them.

ETB (1/7) (End of Transmission Block) — indicates th end of some division of data that the transmission system must make, unrelated to an, division in the format of the logical data itself.

One lists each symbol by name and then shows its action between parentheses. The other asks you to depress in turn all the funny keys on your terminal . . .

STICKS 0, 1 — Physical Communication (4)

This group is used for communications. It includes:

NUL (0/0) (Null) — the standard says that it is "used" to accomplish media fill or time fill" . . . "may be inserted into or removed from a stream of data without affecting the information content of that stream." And that's exactly what the standard also says about DELete (7/15), which it lists as a control character even though it is not in the control sticks! The only difference I can see between them is that on perforated tape you can make any character into a DELete, but none into a Null.

CAN (1/8) (Cancel) — the receiver is to disregard the data received up to that point, starting fror restart point that receiver and sender hav agreed upon. It is common in timesharing fo. Cancel (often generated by a Control-X) to work on a line-at-a-time basis, to delete an unwanted string of entry characters, and effectively put one back to the position of re-entering the entire line. In this case, the agreement between sender and receiver is "back to the last CR." But there are many other ways that Cancel could be used, and for parallel as well as serial transmission.

SUB (1/10) (Substitute) — a character that says probably we would have had another character in this position if we could have figured out what it was supposed to be! There are many reasons for such confusion — perhaps parity didn't check out. But it is better to put in a SUB to keep the field lengths and such correct. Moreover, note its symbol, a mirror image (not the Spanish inverted) question mark. If this is displayable, it will tell you definitively that the system doesn't know what it is, and you can make a good guess in many cases, particularly in word text.

EM (1/9) (End of Medium) — defines the previous character as the last usable character on that medium, whether or not there is more recordable space on the medium.

STICKS 0, 1 — Device Control (11)

This group is used for control of devices such as terminals.

HT (0/9) (Horizontal Tabulation) — the standard says that is "advances" the active position to the next predetermined character position on the same line." There are two ways this can work:

1. Right at the terminal, if it has the horizontal tab capability built in. Sometimes you can set the tab positions by using the terminal only; almost always the computer can be made to set the tabs on the terminal. Then when you hit HT during entry, or HT is read from the computer output, the printing or displaying (active) position will skip to the next tab setting.
2. By a formatting program in the computer, which must be given some indication of the tab setting positions in force at any particular point in the file. The program then simulates horizontal tab movement by filling the lines with spaces as needed to achieve the alignment.

VT (0/11) (Vertical Tabulation) — the standard says that it "advances the active position to the same character position on the next predetermined line." And if you agree with somebody else, it can be to the first position in that line instead. This is a very dangerous character to use. It cannot be used directly on any terminal that I know of. Even if it could, the implementation rules are not supplied unambiguously in the ASCII standard. And for use by a formatting program, one would have to predefine the number of lines to be skipped. That's pretty tough when you are inserting and deleting lines, as every programmer knows.

LF (0/10) (Line Feed) — like vertical tab, but just to the next line, which is clean enough. If receiver and sender agree (again as in vertical tab), it can be to the first position of the next line, in which case it is called New Line (NL). Some manufacturers implement this. I personally prefer having a separate Carriage Return and Line Feed. Both codes can be generated with a single keystroke, and they often are.

FF (0/12) (Form Feed) — again like vertical tab, to the same character position unless sender and receiver agree that it is to the first position in the new line, except that the tab is to a new line position that is related to a form of some size (those that fold 11 inches apart, for example). This control could run wild if your terminal or other display device is not equipped to handle it, so use it with caution in files.

CR (0/13) (Carriage Return) — moves the active position to the first position on the *same* line! Not like typewriters. They have effectively incorporated the New Line feature. But the non-advancing CR is better for terminals, even if it is misnamed. Neither video terminals nor ball and daisy wheel typewriters have carriages, so live with it.

BS (0/8) (Backspace) — Backspace is a very tricky character. On some terminals, such as video terminals, there is no key to generate Backspace for entry into the text stream or buffer. On many it can be created via Control-H. Even then, it may or may not be operative.

Backspace is meant for physical movement of the active position (which may or may not coincide with a cursor position, when such exists). Historically, it was included for hardcopy terminals and other hardcopy devices for some of these uses:

- Underscoring (underlining).
- Other forms of highlighting, such as bold. For example, the sequence A BS A BS A would strike the A three times on a hardcopy device, and make it look boldface (such a sequence can also be translated to call a boldface font in photocomposition).
- Editing indications. For example, in legislative bill drafting to indicate the deleted or changed portion:
  This is obsolete.
- Forming comp
  ± ≠ ↑ ↓ ...arian forint)
- Forming accented letters, primarily for European languages. Examples:
  Ä Å Ø (Scandinavian letters following Z)
  Ñ à â ô ü

Warning: Backspace is entirely different from a cursor movement on a video terminal! When the cursor is moved to a position where a character is already entered, succeeding entry in that position usually destroys the original character and replaces it with the new entry.

I personally haven't seen any video terminals with a true backspace. A former president of Infoton told me it could be done as an engineering special for about $5,000 one-time cost.

Warning: There are three ways to create underscored text for hardcopy terminals:

1. The characters, that many backspaces, and that many underscores (or vice versa).
2. A character, BS, underscore, the next character, etc. This is called the canonical form, and is used quite commonly.
3. Underscore, BS, character, underscore, etc.

I have noticed a lot of difficulty moving back and forth between hardcopy (at my home) and video (in my office) terminals. One tends to underscore on the hardcopy terminal and forget that half of the pairs are going to be wiped out by the cursor on the video terminal. In the first two methods above, it's the text that gets wiped out, and it's hard to read on the fly. So if you plan to display a file on a video terminal, find another highlighting method, or use the third underscoring convention. Even that may give problems if done by embedding an underscoring command in the file you pass to a formatting program; most such programs put the underscore last instead of first.

BEL (0/7) (Bell) — sounds an audible signal to get the user's attention. Some terminals are not so equipped, but they should be. It's good human engineering. But please give me an adjustable volume control!

And then there are the four device controls for unspecified purposes, DC1, DC2, DC3, and DC4 — in positions 1/1 through 1/4. Different manufacturers treat these like a wild card in poker — they make them anything that they want. Doesn't lead to much compatibility, so beware.

STICKS 0, 1 — Field Separators (4)

This group is used for formatting and string processing. These are the separators in positions 1/12 to 1/15. I got the idea originally from the Word Mark in the IBM 1401, which used an extra bit in the low-order character in a field as a delimiter. ASCII uses special and separate characters to indicate a hierarchical structure. Originally I put in eight such characters, but only these four remain:

FS(FileSeparator — 1/12)
GS(GroupSeparator — 1/13)
RS (Record Separator — 1/14)
US(UnitSeparator — 1/15)

FS is most inclusive, US the least inclusive. And we can consider the blank/space as the next lower order separator from these. Suppose we had a line of text like this:

(text1)US(text2)US(text3)RS(text4)US(text5)GS(text6)

On many terminals these delimiting control characters would not print, so we would see only a continuous stream. On others they might show as spaces. A TEX command to break the line at the record separator would be:

scan:line:*rs

The variable *left would contain "(text1) . . . (text3)". The variable *right would contain "(text4) . . . (text6)".

STICKS 0,1 — Changing Sets (3)

This group is used for moving to and from alternate graphic and control sets. This includes ESCape (1/11), Shift Out (0/14), and Shift In (0/15).

These basic control characters have permitted design of a quite marvelous structure for extension and expansion. It allows us to code and classify most of the world's graphic symbols for computer storage, interchange, and display. This big area will form most of Part III of this article.

IN THE NEXT INSTALLMENT

The ASCII Collating Sequence
ASCII and Programming Languages
ASCII and Media
Keyboards
ASCII and Display/Printing
Code Extension — Alternate Controls
Code Extension — Alternate Graphics
ASCII and Non-Latin Alphabets
Code Expansion — 8-bit ASCII

WHERE TO GET MORE INFORMATION

There are four sets of Information Processing Standards that may be of concern to you:

- ISO. Sold only through ANSI (American National Standards Institute), which has the franchise. That makes the prices high — much higher than in other countries.
- ANSI. These are American National Standards developed via the X3 and X4 committees, mostly. Prices still pretty high.
- ECMA (European Computer Manufacturers Association), 114 Rue du Rhone, 1204 Geneva, Switzerland). Free, and they have a lot more advanced standards than ISO and ANSI. But a modest donation would not be unwelcome.
- Your friendly U.S. Government, in the person of the Department of Commerce, National Bureau of Standards, Institute for Computer Sciences and Technology, in Gaithersburg, MD 20760. If by any chance you are employed by the U.S. Government, you get FIPS PUBS (Federal Information Processing Standards Publications) for cheap. Otherwise, see ANSI. (Refer to Tables 1a, 1b, and 1c). In many cases they are essentially reprints of the ANSI standards, for a fraction of the cost.

If you can't wait for the standards to be approved and published, catch them in progress. Ask CBEMA, the sponsor of ANSI X3, to put you on an observer list for the committee in your area of interest. The address is:

Robert Brown, Director of Standards
Computer & Business Equipment Manufacturers Association
1828 L Street NW
Washington, D.C. 20036
(202) 466-2288 Telex 89 29042

REFERENCES

1. ANS X3.4-1977, available from the American National Standards Institute, 1430 Broadway, New York, NY 10018.
2. ISO 646, available from ANSI (Reference 1).
3. R.W. Bemer, "ASCII — the data alphabet that will endure, in *Management of data elements in information processing*, National Bureau of Standards, 1975 October, 17-22.
4. R.W. Bemer, "A view of the history of the ISO character code," Honeywell Computer J. 6, No. 4, 1972, 274-282.
5. E.H. Clamons, "Character codes: who needs them?", Honeywell Computer J. 5, No. 3, 1971, 143-146.
6. The TEX Subsystem of the Timesharing System, Series 60 Level 66, Honeywell Information Systems, 200 Smith Street, Waltham, MA 02154, Order DF72.

ACKNOWLEDGEMENTS

Thanks go to co-workers at Honeywell Information Systems: Eric Clamons for much background, insight, and experience gained from working for a long time as chairman of X3J2 — the committee charged with the development of ASCII. And to Pat Skelly, ACM representative on ANSI X3, for collecting all the various national and international standards documentation upon which many of the figures were based.

FOOTNOTES

[1]For those curious about the reverse slash, it came from ALGOL 58. The reference language specified ∧ and ∨ as the symbols for AND and OR respectively. I put the reverse slash in so these could be made as 2-character groups — /\ and \/

[2]You will still see many terminals where this vertical bar is broken in the middle. This resulted from a hassle with the PL/I people, who wanted to stylize the exclamation point (2/1) as a vertical bar for OR in that language. And of course that would make the graphics the same. The compromise (at horrendous cost in people time) was to break the real vertical bar in ASCII. But it turned out that the PL/I people didn't really need it, or else it gained no momentum, so the real vertical bar is back to normal in ASCII-1977. Let's fix those terminals.

[3]The Italians also have a different solution to hyphenation and right justification. It ignores the syllable structure and simply demands that if, when you get to the last position in the line, the current word is not yet completed, that last character shall be underscored, and the word continued without fuss on the next line. I rather like it.

THE FATHER OF ASCII, Robert W. Bemer

Robert "Bob" Bemer received his A.B. in Mathematics from Albion College in 1940, and a Certificate in Aeronautical Engineering from Curtiss-Wright Tech a year later.

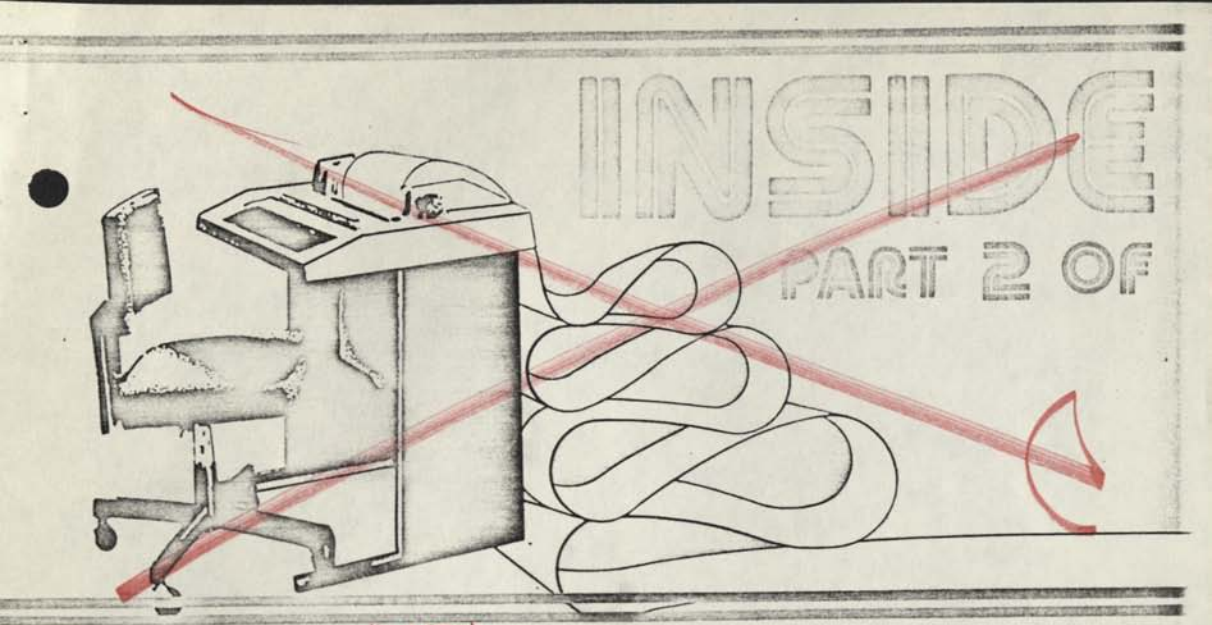
His vast work experience includes employment with the major leaders of the aircraft and electronics industries—most recently, as Senior Consulting Engineer with Honeywell Information Systems.

Highlights of his many accomplishments include: The discovery of polynomial telescoping (1954); creation of the PRINT 1 programming system for IBM (1956); development of FORTRANSIT; development of COMTRAN, one of the three major inputs to COBOL; development of XTRAN, predecessor to ALGOL (1958); was a major influence in the choice of the 8-bit character in IBM System 360 (1960); an influence in building the 1108 and 6000 systems; and editor of Honeywell Computer Journal.

He has an impressive list of over 71 publications to his credit.□

INSIDE ASCII

PART 2 OF



The abstract aspects of ASCII were treated in Part 1. Now we come to some aspects of usage and implementation. Certainly one major use area is the ordering of files.

THE ASCII COLLATING SEQUENCE

To put items in some ordering, the entire precedence relationship for that ordering must be defined. Higher or lower, precedes or follows, or whatever. For single characters, this ordering relationship is called the "collating sequence".

The ASCII standard used to say that the collating sequence for both graphics and control characters is defined simply by their binary representations. Later it added a warning that this collating sequence "cannot be used in many specific applications that define their own sequence". What an understatement!

The 1977 version hedges and speaks all around the problem without making it clear. It's not all that difficult. Suppose you have two files, and you want to know how they differ and/or how they are the same. For this purpose, the implied collating sequence (straight binary comparison) is just fine. The two files will be in the same order, and can be matched.

Whether that straight binary ordering can be used for any other purpose is doubtful. It won't work for signed numbers.

Ordering Numerals

Take these four values: 22, 13, minus 6, and minus 31. If the sign is placed before the digits, ordering by the ASCII collating sequence yields:

+13
+22
−06
−31

This is obviously worthless. It's because ordering is decided left to right, and the minus sign has a binary value 2 higher than the plus sign. Or if the sign were to follow the numeric values we would get:

06−
13+
22+
31−

because the complete decision is made in the leading digit. Again, a worthless sequence.

The way to achieve a proper ascending sequence is to separate the values into two groups, ordering those with plus signs in ascending sequence, and those with minus signs in descending sequence. Then put the plus group following the minus group. And vice versa for a total descending sequence. Notice that this works regardless of whether the sign precedes or follows the digits.

Ordering Alphabetic Fields

Alphabetic ordering is even more complex, particularly in handling both upper and lower case. Again the implied ASCII collating sequence can go wrong. People who have not studied the collating problem for data containing both upper and lower case are inclined to jump to wrong conclusions. I did myself, for the IBM Stretch computer in 1958, assigning the ascending binary sequence as AaBbCc. Using this for a telephone directory would give us the lefthand column. The straight binary sequence of ASCII would yield the righthand column, just slightly different:

De Carlo	De Carlo
De La Rue	De La Rue
De Long	De Long
DeLair	DeLaRue
DeLancey	DeLair
DeLaRue	DeLancey
Delancey	Delancey
de Carlo	de Carlo
de la Rue	de la Rue
deLancey	deLancey

Either version will get a lot of anguished subscribers!

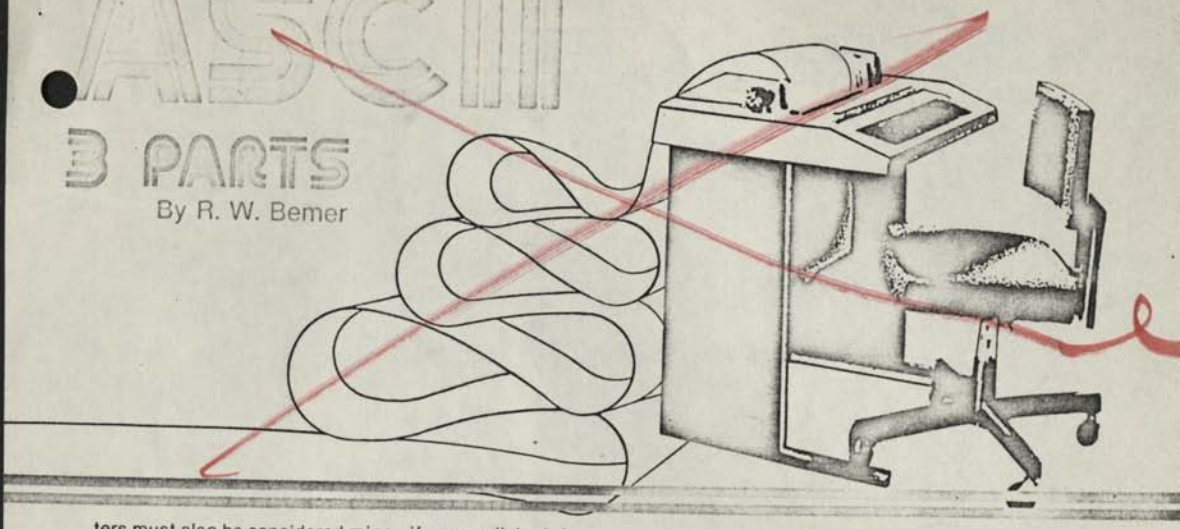
In the simplest case, two alphabetic items must be compared with the case ignored. Only if they are *then* equal is case called into consideration to break the tie, and it is also applied successively left-to-right!

In short, the upper and lower case versions of a letter do not both get full graphic significance. Typing either "Y" or "y" will indicate a "yes" reply, but "N" will not. Because the case distinction is minor, comparisons must first be made on major distinctions, with the minor distinctions used only as tie-breakers. Accenting of let-

ASCII

3 PARTS

By R. W. Bemer



ters must also be considered minor, if accomplished via backspace, but this leads us into rules controlled by foreign governments, and won't be considered here.

Real life is more complicated than this. The ordering and sequencing of characters and words cannot always be accomplished by simple binary comparison of codes. There are constructions such as O'Reilly, l'Informatique (as data processing is called in French), and Smith-Jones — to say nothing of the Juniors, IIIs, Esq., FBCS (which I am), and so on.

Making an ASCII comparison, with the case as a minor, gives us:

De Carlo
de Carlo
De La Rue
de la Rue
De Long
DeLair
DeLancey
Delancey
deLancey
DeLaRue

Because we at first ignored case here, De Carlo and de Carlo have identical bit patterns. Tie-breaking is done by appending the binary pattern representing case, "0" for upper, "1" for lower. Specifically, 01001111 for De Carlo, 11001111 for de Carlo.

	D	E		C	A	R	L	O	
De Carlo	44	45	20	43	41	52	4C	4F	(4F)
de Carlo	44	45	20	43	41	52	4C	4F	(CF)

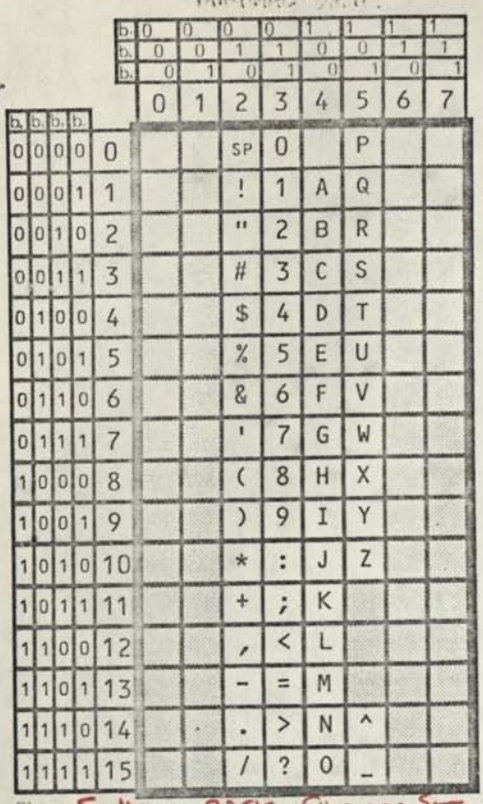
But even this method will not put "DeLaRue" and "De La Rue" in the same cluster. And surely this is desirable and even mandatory. It will require some special handling for spaces. The New York Telephone Company's document on this problem runs to several pages! They'd probably give you a copy upon request. You might need to know those rules before trying one of the toughest acts in data processing — putting last name first, or vice versa.

Using Controls in Ordering

There is one more aspect of ASCII useful to the order-

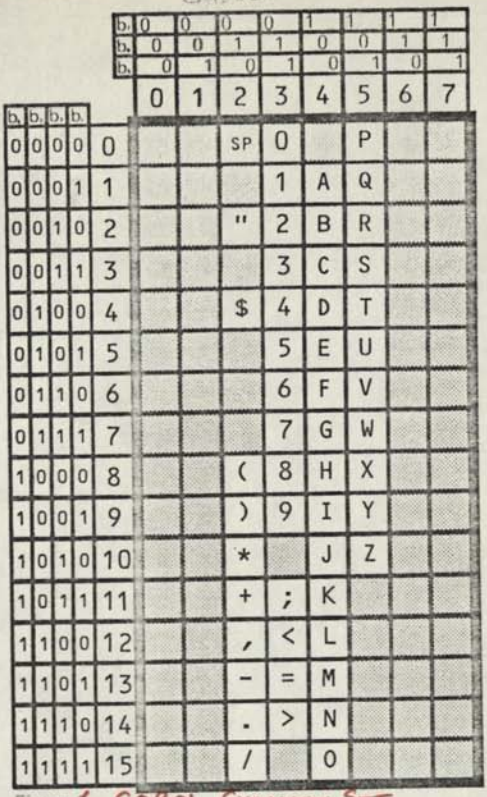
b7	0	0	0	0	1	1	1	1
b6	0	0	1	1	0	0	1	1
b5	0	1	0	1	0	1	0	1
b4 b3 b2 b1	0	1	2	3	4	5	6	7
0 0 0 0 0			SP	0		P	←	→
0 0 0 1 1				1	A	Q	α	∨
0 0 1 0 2			"	2	B	R	⊥	ρ
0 0 1 1 3				3	C	S	∩	r
0 1 0 0 4				4	D	T	L	~
0 1 0 1 5				5	E	U	∈	↑
0 1 1 0 6			&	6	F	V	⌶	U
0 1 1 1 7			'	7	G	W	∇	ω
1 0 0 0 8			(	8	H	X	Δ	⊃
1 0 0 1 9			)	9	I	Y	ι	∧
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. APL CHARACTER SET



b7	b6	b5	b4	0	0	0	0	1	1	1	1
				0	0	1	1	0	0	1	1
				0	1	0	1	0	1	0	1
				0	1	2	3	4	5	6	7
0	0	0	0	0			SP	0		P		
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 5. Minimal BASIC Character Set



b7	b6	b5	b4	0	0	0	0	1	1	1	1
				0	0	1	1	0	0	1	1
				0	1	0	1	0	1	0	1
				0	1	2	3	4	5	6	7
0	0	0	0	0			SP	0		P		
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 6. COBOL Character Set

ing problem. In the days of punch cards, before computers, one often used several card files related by a key. A sorter (with pockets for the cards to drop into) might be used to select the cards for all redheaded females between 18 and 24 years of age. But these cards would have only the employee number and such characteristics on them. To get the name, address, and telephone number one might have to go to a second (related) deck of cards. So the first deck (the subset of interest) would be placed in the first hopper of a collator, and the deck with all names and phone numbers in the second hopper. Then a card would be fed from the first hopper, followed by successive cards from the second hopper, until a match was found on employee number. Obviously both decks had to be in the same ordering for this to work, and thus the term "collating sequence".

In effect, we were sticking the cards of the first deck upright just in front of the corresponding cards of the second. To do this with ASCII requires that we have characters that collate lower than the lowest graphic, the space (2/1). We do have them. The best to use are NUL, FS, GS, RS, and US. Put one of these after each search key, then put the two files together and order them as adjoined. Now those records having a search key with one of our five control characters appended will precede the corresponding record having an ASCII graphic following the key.

Note that the four information separators (FS, GS, RS, US) are designed to collate just behind Space, in that order. This contiguity means that they can be used as a hierarchy of spaces of different class.

Other Collating Features

ASCII was designed when there was substantial investment in files already ordered on a Topsy-class IBM sequence, where the basic punctuation was low to the alphabet, but the digits were high to it. How then to accommodate this and still provide a 4-bit subset? My morning shower provided a solution (it still does!).

The 4-bit subset is formed of the first 10 graphics of stick 3 (the digit graphics) and the last 6 of stick 2. This jog was shown shaded in the early forms of ASCII, but has all but disappeared from memory now. It enables stick 3 (with the digits and new special graphics) to be ordered high to all the others via passive logic, thus overcoming opposition to the adoption of ASCII.

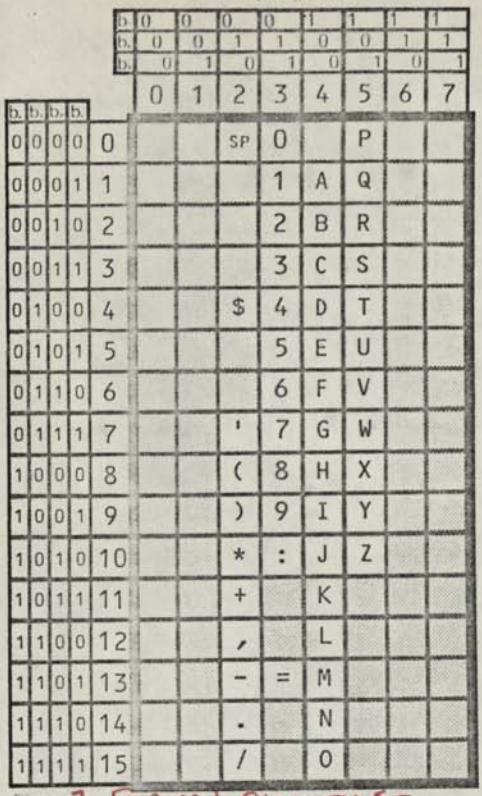
ASCII AND PROGRAMMING LANGUAGES

Standard ECMA-53 (1978 Jan), "Representation of Source Programs for Program Interchange," gives the subsets and/or modifications of ASCII as they are used for these five programming languages (Footnote 1):

Language	NO. OF CHARACTERS USABLE: Subset of ASCII	Other
APL	57	32
Minimal BASIC	60	0
COBOL	51	0
FORTRAN	49	0
PL/I	55	2

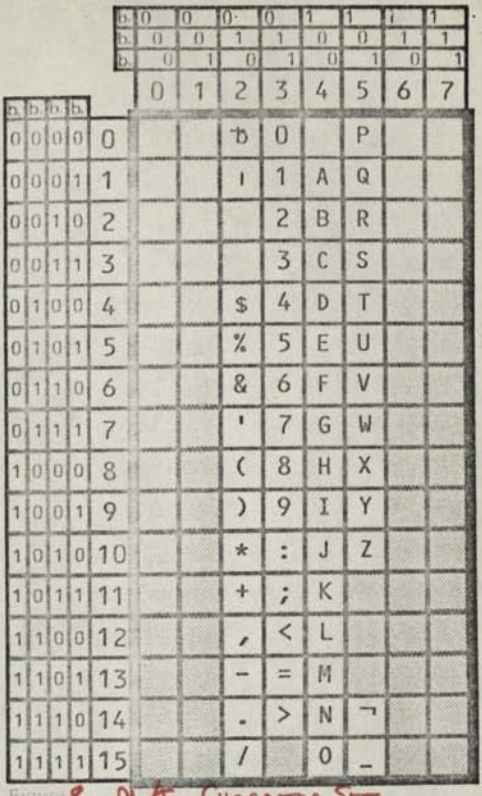
Figures 5 through 6 are the character sets for these languages as given in ECMA-53. They show the only characters permissible for use in source programs, except for:

non-numeric literals	in COBOL
comment-entries	"
comment lines	"
character constants	In FORTRAN
comments	"



b7	b6	b5	b4	0	1	2	3	4	5	6	7
				b3	0	0	0	0	1	1	1	1
				b2	0	0	1	1	0	0	1	1
				b1	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 7. FORTRAN Character Set



b7	b6	b5	b4	0	1	2	3	4	5	6	7
				b3	0	0	0	0	1	1	1	1
				b2	0	0	1	1	0	0	1	1
				b1	0	1	0	1	0	1	0	1
0	0	0	0	0			b	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 8. PL/I Character Set

character-string-constants in PL/I
comments "

For these purposes only, other ASCII characters may be used, providing there is agreement between the sender and receiver for any interchange of source programs.

The TEX language has gone farther than this general caution. There the specific characters have permanent names. For example, one could say:

```
linefeed = "
           "  (actual line feed inside the quotes)
if lf:eqs:linefeed ....
```

and it would be true, because "*lf" is the permanent name of Line Feed. The control characters have names that are the letters from the ASCII chart, preceded by the asterisk to show that they are read-only variables with permanent content. TEX can in fact operate upon all 256 characters of ASCII in an 8-bit byte, all 512 in a 9-bit byte.

Specific Notes on the Figures

APL -Sticks 6 and 7 (ordinarily lower case alphabet) are replaced entirely except for the DELete position.
-Space is nonprinting, although the symbol shown is SP.
-Ampersand (2/6) is not used for writing source programs, except as the last character of a line if that line is to be continued on the next line.

PL/I -In position 2/1, the exclamation point is replaced by a vertical bar for OR.
-In position 5/14, the circumflex is replaced by the symbol shown, for NOT.
-If you have to use your terminal for both PL/I and some other programming language, forget that foolishness. You can get by with the exclamation point as OR, and the circumflex as NOT. The important point in source program interchange is to have the encoded representations of the characters exchanged correctly.

(all) -Although the character BLANK (space) is shown as the flagged lower case "b" in the FORTRAN and PL/I sets, there is no printing graphic to indicate it. For all practical purposes, it is really the Space of ASCII (2/0).
-Four of these five languages (not APL) have the "$" shown in 2/4. When the International Reference Version of the code is used, this becomes the universal currency symbol, which is also acceptable.
-Minimal BASIC uses "#", which is the International Reference Version symbol. The national symbols, such as the English pound sign, are also acceptable.

ASCII AND MEDIA

ASCII and Punch Cards

Reading and punching equipment for punch cards, being very mechanical, is so expensive that microcomputer people are unlikely to use them. So you might ask why we bother here with the representation of ASCII on this medium? I can think of at least three reasons:

- A scientist at the U.S. National Bureau of Standards said once that if punch cards were on the way out, it was the only product he ever saw dying on an upward usage curve. Thus they are likely to be around for a long

	ISO	ECMA	ANSI	FIPS PUB	CSA	BS	AS	CCITT	JIS	GOST
Hollerith Punched Card Code	1679 2021	44	X3.26-1970 $4.25	14	Z243.14 .36	4636/3 /4	1063			
Track Assignment - 25.4 mm Perf. Tape	1113	10	X3.6-1965 $3.00	2	Z243.8	3880/3	1062		C6221	
Track Assignment - 12.7 mm Mag Tape 200 cpi NRZI 9-track	1862	5	X3.14-1973 $3.25			3968	1008			
Track Assignment - 12.7 mm Mag Tape 800 cpi NRZI 9-track	962 1863	12	X3.22-1973 $3.75	3-1		4503/1	1009		C6222	
Track Assignment - 12.7 mm Mag Tape 1600 cpi PE 9-track	3788	36	X3.39-1973 $3.75	25		4503/2				
Track Assignment - 12.7 mm Mag Tape 6250 cpi GCR 9-track	DP 5652		X3.54-1976 $5.25	50						
Labeling & File Structure - 12.7 mm MT	1001	13	X3.27-1977 (unpriced)		Z243.7	4732	1068			
Track Assignment - Magtape Cassette 3.81 mm, 32 bpmm	3275 3407	34	X3.48-1977 $5.75	51		5079/1				
Labeling & File Struct. - 3.81 Magtape Cassette	DIS 4341	41								
Track Assignment - 6.35 mm Cartridge Tape 64 bpmm PE	DIS 4057	46	X3.56-1977 $4.24							

Table 1a. Standards for ASCII on Physical Media.

	ISO	ECMA	ANSI	FIPS PUB	CSA	BS	AS	CCITT	JIS	GOST
Bit Sequencing in Serial Transmission			X3.15-1976 $3.00	16-1				V.4 X.4		
Char. Structure & Parity Sense - Serial-by-Bit			X3.16-1976 $3.50	17-1				V.4 X.4		
Char. Structure & Parity Sense - Parallel-by-Bit			X3.25-1976 $3.50	18-1				V.4 X.4		
Procedures for Using Commun. Control Chars.	1745	16	X3.28-1976 $10.50		Z243.13	4505/1	1484/1			
Message Heading Formats	1745		X3.57-1977 $5.25							
Advanced Data Commun. Control Procedures			BSR X3.66							

Table 1b. Standards for ASCII in Communications.

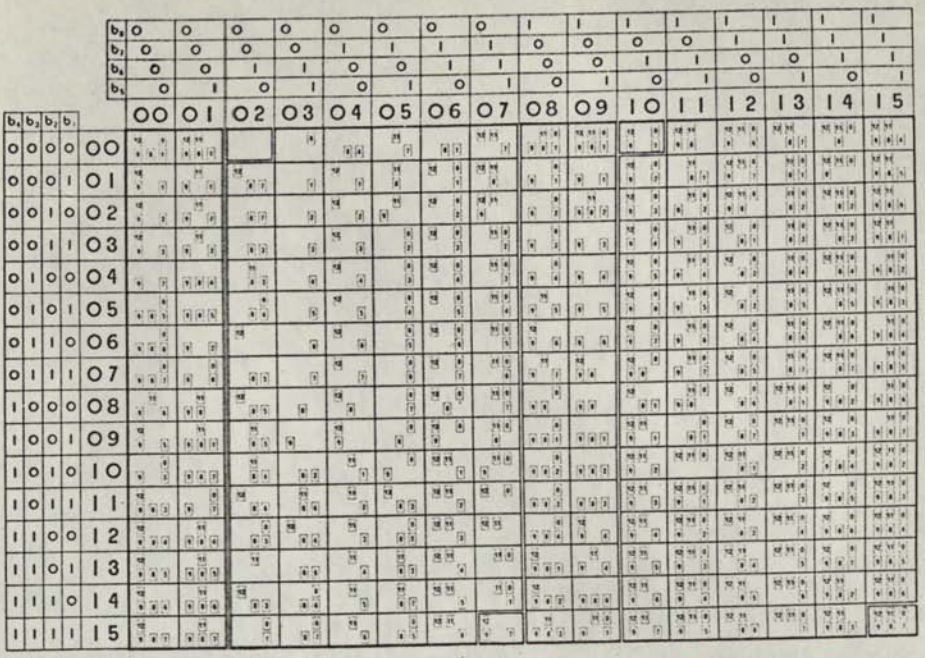


Figure 6. CARD HOLE PATTERNS ASSIGNED BY CODES

FIGURE 9. CARD HOLE PATTERNS ASSIGNED BY CODES

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

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

Figure 7. CODES ASSIGNED BY CARD HOLE PATTERNS

FIGURE 10. CODES ASSIGNED BY CARD HOLE PATTERNS

time, and you may need to transfer some of those files to other media that you do use.

- There is some likelihood that microcomputers could be used in the reading and punching equipment itself, to make it less expensive.
- ASCII users are going to be confronted for a while yet with one of the several versions of IBM's EBCDIC, and the punch card assignments provide the only legitimate link for conversion of EBCDIC files to ASCII.

So Figure 4 defines the hole patterns for the binary encodings. And Figure 5 defines the encodings for the hole patterns. Don't worry about the inconsistency in the relationships. Nothing can be done about it now, because it started with Herman Hollerith's first U.S. Census machines in 1890. At first only digits and + and − signs were used. Then the code was expanded to the upper case alphabet. And other special characters for commercial use. When FORTRAN came along in 1964, it turned out that the limited capability of the subset of a 6-bit set would not permit the graphics needed for scientific work. For a long while there were dual graphic representations for several of the punch card code combinations, and this carried over into printer chains, and so on.

The only logic that the patterns follow is that they do or do not have a punch from among these six possibilities:

12-punch (top row)
11-punch (next to the top row)
0-punch
8-punch
9-punch (bottom row)
a punch from among the digits 1 through 7

Including the no-punch-at-all combination (NUL), this gives 256 combinations, just right for the 8-bit code. Although ASCII was technically only a 7-bit code at the time this rule was formulated, it was felt necessary to plan ahead a little.

ASCII and Magnetic Tape

Figure 6 gives a compact representation of several relationships, among which is the assignment of ASCII bit pattern to 9-track magnetic tape. The jumbled assignment may remind you of the "firing order" for the cylinders of an automobile engine. In fact, we used to call it just that. It was intentional for increased reliability. As in so many cases, better technology has removed the need for peculiar design, but the assignments are unchangeable because of data file investment.

There is no parallelism in recording and reading on cassettes and cartridges. The ASCII bits are recorded serially in the track. Thus Figure 6 does not consider these media.

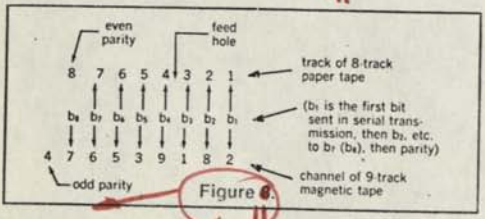


Figure 6.

ASCII and Communications

Not only is the topic of ASCII and communications a very complex and large dissertation for this article — it is also undergoing substantial rethinking, enlargement, and invention. You will have to follow on your own the workings of the CCITT, the various networking systems of the several large and many small manufacturers of computer systems, and the offerings of the common carriers — either on the local distribution system (via ATT) or direct distribution (via Satellite Business Systems).

Many of the existing standards are listed in Table 1b. Many more are under development. Arguments are raging internationally on the merits of packet switching, byte protocols, value-added systems, open-working systems, tariffs, data movement across national borders, the X.25 protocol, etc., etc. ATT is offering a new service because they suddenly discovered data-undervoice (DUV). All I can tell you now is that it is all based upon ASCII, and the proposed protocols are all dependent upon the ASCII control characters in stick 0 and 1. It will take years for this to shake out, and for now all one can do is get on the CBEMA mailing list (see reference, Part I, INTERFACE AGE, May, 1978).

ASCII AND THE METRIC SYSTEM

The full ASCII graphic set (both cases) is sufficient to indicate all symbols and prefixes of the SI (International System of Units, the new metric system), with three exceptions. They are the Greek letters "omega" for "ohm", and "mu" for "micro", and the degree symbol for Celsius temperature. These three characters will be provided in 8-bit ASCII (see Part III, next month). Meanwhile, for these, and also for such equipment that has only a single case, there is a standard way of representing the SI units and prefixes. This is given in International Standard 2955, "Representations of SI Units and Other Units for Use in Systems with Limited Character Sets", and also in American Standard X3.50-1976.

To keep the record straight, let's first look at the characters used for the prefixes. They're shown in Table 2, which indicates multiples from 10 to the − 18 up to 10 to the + 18:

10 + i	i	10 − i
exa (E)	18	atto (a)
peta (P)	15	femto (f)
tera (T)	12	pico (p)
giga (G)	9	nano (n)
mega (M)	6	micro (μ)
kilo (k)	3	milli (m)
hecto (h)	2	centi (c)
deka (da)	1	deci (d)

Table 2. Metric Prefixes

Above 3 there are no powers except multiples of 3. This practice breeds better comprehension, like marking off three's in writing numbers of many digits. Also, as a memory convenience, all symbols are upper case for powers greater than + 3. And there are no conflicts with the symbols for the units of measurement.

Now, again for the record, here are the ASCII character(s) used as symbols for the units:

A	ampere	cd	candela
Bq	becquerel	d	day
C	coulomb	g	gram
°C	degree celsius	h	hour
F	farad	l	litre
Gy	gray	lm	lumen
H	henry	lx	lux
J	joule	μ	micro
K	kelvin	m	metre
N	newton	min	minute (time)
Ω	ohm	mol	mole
Pa	pascal	rad	radian
S	siemens	s	second (time)
T	tesla	sr	steradian
V	volt	t	tonne/metric ton/ megagram
W	watt		
Wb	weber		

Table 3. Metric Units

Table 3 shows the rules clearly. Units not named after people are all lower case, as shown in the righthand column (although I do know a Mr. Day). In the lefthand col-

umn are the units that are named after people. The names of the units are not capitalized at all, but the symbols begin with an upper case letter.

I said previously that there were no conflicts between unit and prefix symbols. But you've probably noticed "d" for both "day" and "deci", "h" for both "hour" and "hecto", "m" for both "metre" and "milli", and "T" for both "tesla" and "tera". OK. But there isn't any confusion in actual usage, because the prefix precedes the unit:

dd	is a deciday (2.4 hours)
hh	is a hectohour (100 hours)
hH	is a hectohenry (but don't ever use the term)
mm	is a millimetre
Mm	is a megametre (1/300 the speed of light)
TT	is a teratesla (Wow!)

I am not suggesting that the prefixes should be applied to other than the primary metric units (the second is the primary time unit; hour and day are not), even though the timesharing system I customarily use figures my time in millihours. But when you get accustomed, the prefixes are very valuable in other ways. For example, an American billion is a kilomillion, whereas the British billion is a megamillion! And my metric teaching program understands such things as kilofathoms.

The "space" character is also vital to correct SI usage. It must occur between values and units, like 123.6 mm, and 22 °C.

And don't forget another peculiarity of ASCII as an international alphabet: (1/14) is absolutely not defined as a "decimal point" (nor is it defined as "period", which in the United Kingdom is "full stop"). For most of the rest of the world, the comma (1/12) is the decimal marker, and the period is used to mark off threes. That's why the recommended practice for marking off threes is to use the space, not either comma or period. E.g., "1 234 567 mm".

To save you the bother of looking up the standards for use with limited character sets, here is the algorithm:

1. If you have ASCII with both cases of alphabet, the three missing symbols are handled as:

ohm	for Ω
Cel (initial cap)	for °C
u (lower case)	for μ (micro)

2. If you have only one case of alphabet (either upper or lower), use it, and these three replacements remain as:

OHM		ohm
CEL	or	cel
U		u

And in addition:

S (siemens)		SIE	sie
h (hour)	become	HR	or hr
t (tonne)		TNE	tne

Examples:

16 UOHM is 16 μΩ

373.15 K = 100 Cel

Notice that no plurals are used in symbol combinations — MICROOHMS, but UOHM.

ASCII AND KEYBOARDS

Technically, a keyboard is an ASCII keyboard if it generates the proper codes for the full set of ASCII graphic and control characters. Moreover, none of the graphic characters should have any control properties.

There are many types of special keyboards — Dvorak, a two-sided one used like an accordion with the hands in a vertical plane, Touch-Tone and its derivatives, etc. There are no formal standards to relate these keyboards to ASCII. For typewriter-style keyboards, however, there are two versions given in the American National Standard. One is derived from the usual electric typewriter keyboard, the other is called the "bit-paired" keyboard. Only the bit-paired keyboard will be shown and discussed here, because the other form is the subject of proposals for extensive change due to the growth of Word Processing. ANSI Committee X4A12 is studying this now.

The bit-paired keyboard was designed for minimum circuitry cost. Thus the "at" symbol (4/0) is paired with the accent grave (6/0), "A" (4/1) with "a" (6/1), and "+" (2/11) with ";" (3/11). Thus the shift key affects each other key by only a 1-bit change.

This keyboard is shown in Figure 8. It is the interchange keyboard of ECMA-23. The numbered arrows key to the notes on changes that would make this ECMA keyboard into the ANSI keyboard for ASCII. It is also equivalent to the keyboard of ISO Standard 2530-1975 (Footnote 2).

Notes for Figure 8

1. *For the ANSI keyboard, this key is put to the right of the circumflex key, on the top row (see Note 6). The Shift Key is put in its place.*
2. *If this key exists and is available, the ECMA and ISO standards put the underscore here, removing it from the "zero" key.*
3. *The ANSI keyboard of course puts a "$" here in place of the international currency symbol.*
4. *This is where the underscore is removed for the 48-key keyboard (see Note 2).*

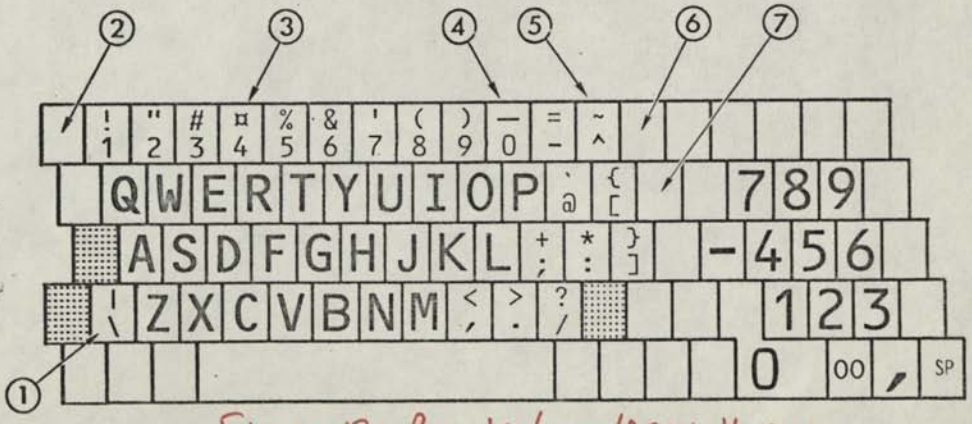
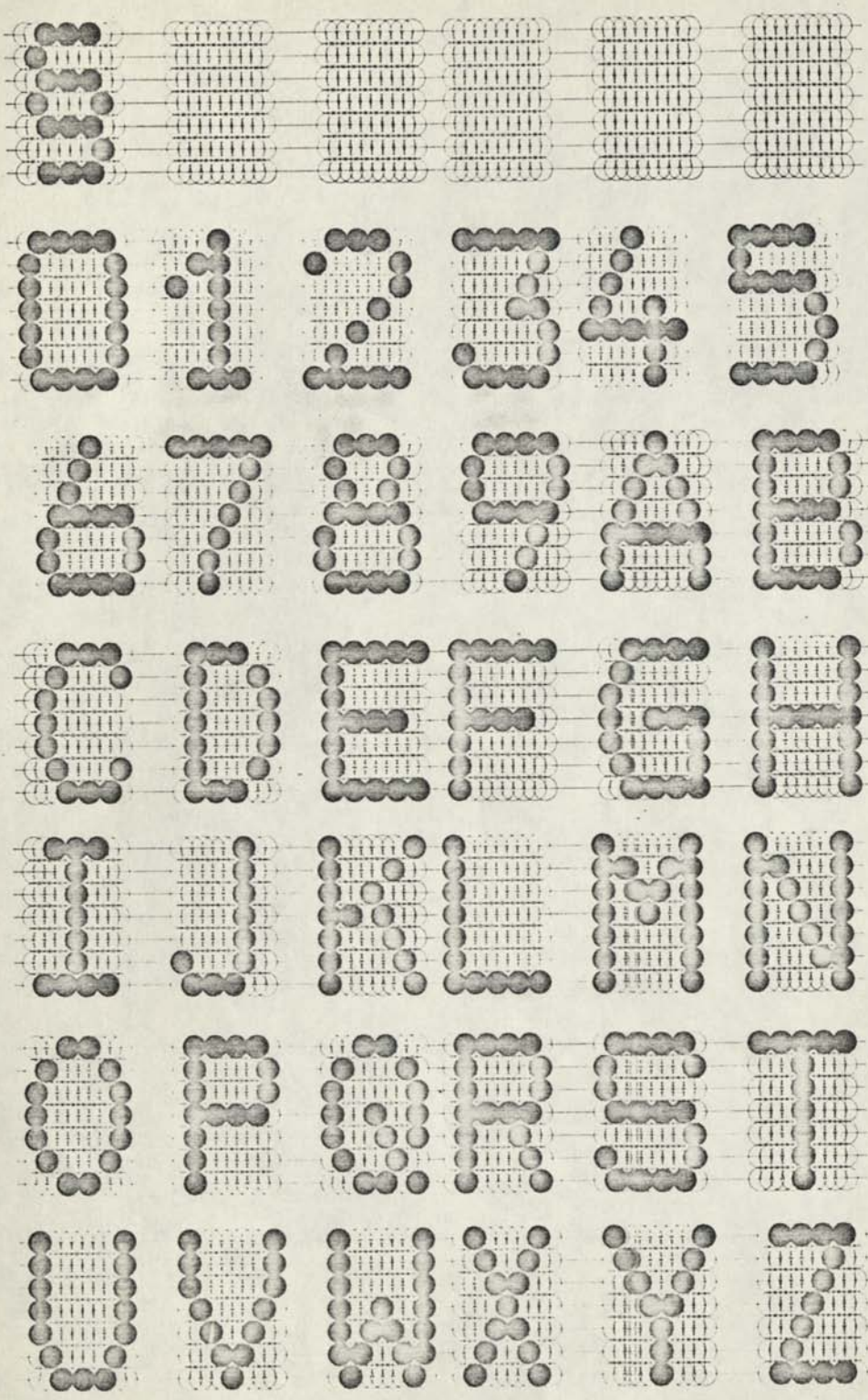


FIGURE 12. BASIC ISO/ECMA/ASCII KEYBOARD



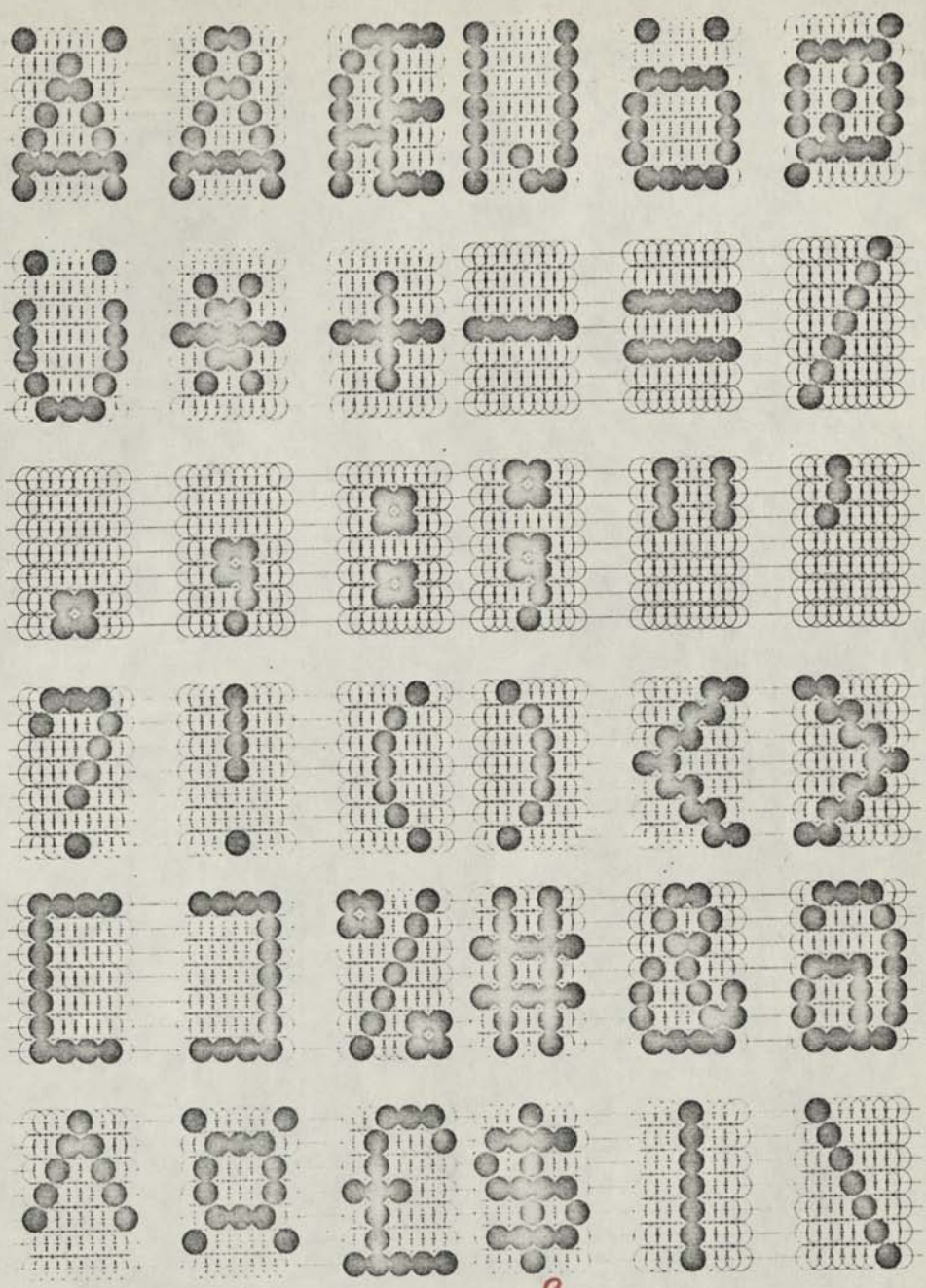


Figure 10.

FIGURE 13. 7×9 DOT MATRIX SHAPES

5. *Here ECMA and ISO show the "overline" instead of the "tilde." It's a question of styling.*
6. *The ANSI keyboard has the reverse slash and vertical bar here, rather than between the shift key and "Z" (see Note 1).*
7. *The ANSI keyboard specifies the underscore here, in both shifts, rather than the positions shown as options in Notes 2 and 4. Practically no keyboards follow this. In fact, as I am entering this text, this is the only key where my Infoton Vistar deviates from the ANSI standard. It has Line Feed there, with Return to its right — a very sensible arrangement.*

Customarily, the Control Key is also tied to bit-pairing in such keyboards. The standards recommend that characters created in combination with the Control Key should use the graphic key in sticks 4 or 6 units higher. Thus "X" (5/8) or "x" (7/8) in combination with the Control Key produce CAN (1/8). Unfortunately this also means that Control-C generates ETX (0/3). And whereas Control-X as CAN is used frequently, to erase an input line of text, ETX is not often wanted. Yet it is a common miskeying to hit C rather than X. In many timesharing systems you will get a disconnect rather than a line delete.

Control and Function Keys

The so-called "QWERTY" arrangement is prevalent throughout the world. Even the French "AZERTY" set is being considered for change. But on top of these basics there are hundreds of keyboard varieties. Some of them have "dead keys" (i.e., the platen or printing element is not advanced when they are hit). This avoids having to use BS for accented letters, but it also creates difficulties in code generation.

There are some general good practices that ASCII keyboards should follow. To facilitate usage by those experienced with typewriters, all controls not used with typewriters should be located outside the customary touch-typing area. As a specific example, the Break/Interrupt key should be located where it is a definite effort to reach it (not mixed in with the keyboard). ISO 3244 may be consulted for these considerations.

Function Keys are those that generate sequences of more than one ASCII character. Examples are cursor keys, Erase-to-EOL, etc. They should be located in special clusters. Most importantly, they must all generate ASCII codes for transmission when in character-at-a-time mode. I know of video terminals where the cursors do not generate codes, as they should not while in full page buffered mode; but they still operate in line mode without generating codes. In this case the screen is alterable, but there is no way of detecting it in the computer.

Many keyboards will have some function keys that are unlabeled, for do-it-yourself assignment. These should also be clustered separately, and generate code sequences when in line mode.

ASCII AND DISPLAY PRINTING

When ASCII characters are displayed, it may be on a video screen, paper, or COM (microfiche).

On the video screen there are a number of methods to form the characters, mostly at the manufacturer's preference. They are usually at pica (constant-width) spacing for economy, so an approximation of graphic quality (such as typesetting) is not obtainable. When lower case is available, the risers and tails extend above and below the line for some screens. In others, they fall within the boundary lines of the upper case characters. They may be shown in inverse video (light background block), or highlighted by different brightness or blinking. Controls for this work will be taken up in Part III of this article.

For paper copy one usually finds either direct impact of a formed letter, or stylus printing. Either method is suitable to proportional spacing if desired. Recently there has been a general trend toward using the 7x9 dot matrix shapes of ECMA Standard 42 for stylus printers. This set of graphics is shown in Figure 13.

For hard print elements, of course, one can get a nearly infinite variety of styles and fonts. There are only two, however, specifically associated with computers — OCR-A and OCR-B. "OCR" stands for "Optical Character Recognition", meaning that the shapes are so styled that a computer-controlled scanner can read the characters as printed on paper, and encode them directly from their shapes.

OCR-A is not suitable for human reading. It's the funny looking one with the diamond-shaped letter "O". I won't dignify it by showing the font here. It was thought formerly, with technology of that day, that making humans work harder to read letters would make it easier and thus cheaper for computers to read them. This argument turned out to be specious, and with today's technology there is no need to use anything other than OCR-B.

OCR-B is specified in ISO 1073/2, ECMA-30, and ANSI X3.49. It is the font shown in Figure 14. I have it on my IBM golfball typewriter at home, and on my daisywheel element at the office. So it should be available for most hard elements, including the carousel type.

The first six rows correspond to ASCII sticks 2-7. In the first row, the pound and universal currency symbol are for replacement as needed. In the fourth row, the underline is discontinuous; a continuous form is shown in the auxiliary set. This set also contains a matching accent acute instead of single quote, the real circumflex (not an up arrowhead), a cedilla, and an "m" of better proportion. □

```
 ! " # £ ¤ $ % & ' ( ) * + , - . /
 0 1 2 3 4 5 6 7 8 9 : ; < = > ?
 @ A B C D E F G H I J K L M N O
 P Q R S T U V W X Y Z [ \ ] ^ _
 ` a b c d e f g h i j k l m n o
 p q r s t u v w x y z { | } ~
       Ä Å Æ IJ Ñ Ö Ø Ü
       ß æ ij ø ß § ¥
       " ´ ^   m ¯
```

Figure 14. OCR-B Font

Footnotes

1. With the distribution, ECMA said "ECMA-53 is an attempt to improve portability of programs. It links the language character sets defined by the language standards, their coded representatives by means of the 7-bit code and the implementations on data carriers (punched tape, punched cards, magnetic tape and magnetic tape cassettes and cartridges). It is a standard of a new type in which already standardized features are assembled in a new standardized combination aimed at supporting interchange and decreasing implementation dependency."
2. ISO 2530 is for the alphanumeric area of the keyboard only. It is augmented by ISO 3243-1975 — Keyboards for Countries whose Languages have Alphabetic Extenders, Guidelines for Harmonizations, and also by ISO 3244-1974 — Principles Governing the Positioning of Control Keys on Keyboards.
3. The fact that these are "guidelines" and "principles" indicate the complexity of the subject. Typewriter manufacturers now supply over a hundred different keyboard arrangements, as their catalog will indicate.

INSIDE ASCII

CODE EXTENSION — GENERAL PRINCIPLES

Over ten years ago it was recognized that ASCII was the basis for codification of the various symbols used throughout the world. Through it, libraries could store encoded books as well as printed books. And while electronic mail may be quite simple with ASCII and its Roman alphabet, that's not the alphabet of all countries. The USSR uses Cyrillic, the Japanese use Katakana, and the Arab world uses its own semi-script alphabet. Moreover, to send a mathematics textbook by electronic mail one would have to be able to encode the formulas and special symbols peculiar to mathematics, which includes many Greek characters!

This is where the ESCape character and ESCape sequences come in. You can get the whole complicated story from ISO Standard 2022 (or ECMA-35) on Code Extension Procedures. But it will be easier to think of reproducing many ASCII Code Tables on the pages of a book, then replacing the ASCII symbols on all but the first page with the other alphabets we need.

Then we make sure that everyone in the world has the same (code) book. (The resemblance to military code books is intentional.) That's done by registering the page number assignment to characters (actually either a control set or a graphic set, but not both) with the French Standards Body AFNOR. That's the Association Francaise de Normalisation, Tour Europe Cedex 7, 92080 Paris La Defense, FRANCE. But you'll find it perhaps easier to get it from ANSI (see data in the first installment, INTERFACE AGE, May 1978).

The registration procedure is spelled out in ISO Standard 2375. It is carefully controlled to prevent frivolity and cluttering up the assignment books, for that all costs money. But the important control and graphic sets of the world may be registered and assigned their own unique ESCape sequence for calling or invoking them.

CODE EXTENSION — BASIC RULES

The control ESC, when encountered in a datastream, means that all characters following it, up to and including the first character from sticks 3 to 7, have special interpretation. The delimiting character is called a "final" (F). Those between ESC and the final are called "intermediates" (I). All of the codes in stick 2 can serve as intermediate characters in ESCape sequences of 3 or more characters in length. The entire group of characters from ESC through the final is called an ESCape sequence.

ESCape sequences obviously require buffers for interpretation, for we cannot know, when they begin, how long they will be. Sequences of length 2 are for single controls. If the character following ESC is from stick 3, the sequences are for private usage, of the class Fp. If it is from sticks 4 or 5, they mean single controls, of the class Fe, from an appropriate set of 32. If from sticks 6 or 7 (except 7/15), they are of the class Fs, composed of single controls. This is elementary extension.

A more complex type of extension is the simulation of one or more 8-bit character sets by alternating between two 7-bit sets. The home base set consists of the C0 (32 controls) set and the G0 (94 graphics plus space and DEL). The alternate sets consist of the C1 (32 controls) set and and the G1 (94 graphics plus space and DEL). The 8-bit set (it doesn't have to be just theoretical if you have a full 8-bit capability) consists of the four parts C0-G0-C1-G1.

These four types of sets are all invoked (designated) by 3-character ESCape sequences in this manner, where F is the final (3rd) character:

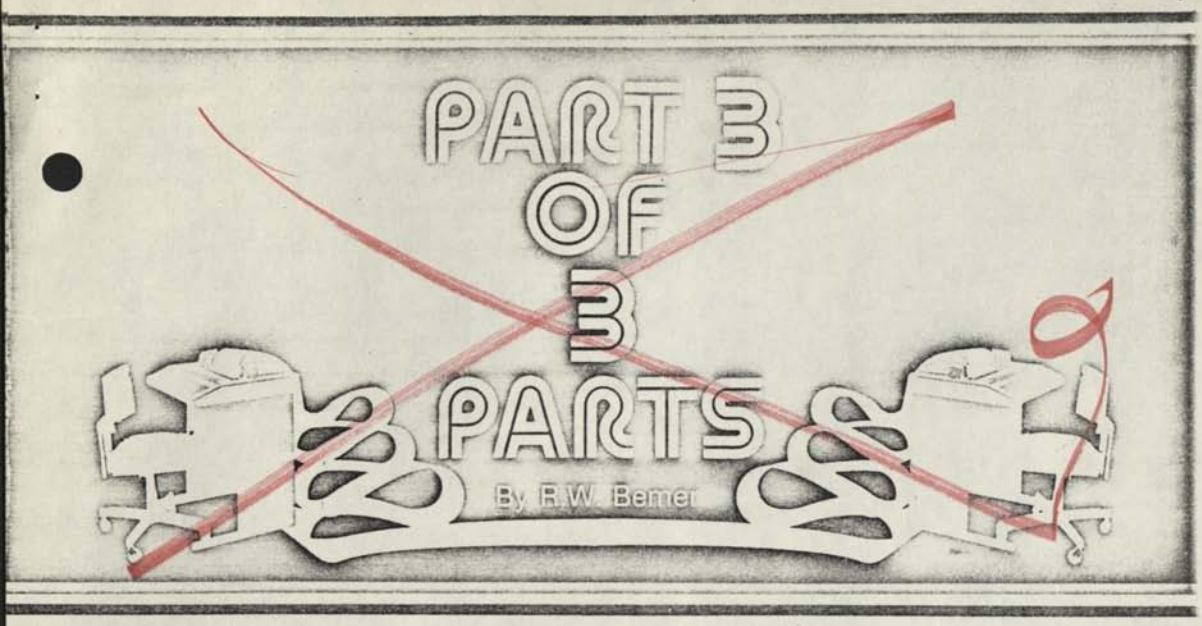
Sequence	Invokes Set Type
ESC 2/1 F	C0
ESC 2/2 F	C1
ESC 2/8 (or 2/12) F	G0
ESC 2/9 (or 2/15) F	G1

The final character "F" selects the particular set to invoke. Once invoked, encountering or entering an SO shifts to the G1 set in force; an SI shifts to the G0 set in force. SO and SI do not affect the control set.

ISO Standard 2022 defines these matters in far more detail, but that is enough for here. That document is complicated and ingenious, and deserves substantial study.

THE CODE EXTENSION REGISTRY

Table 1 identifies the graphic sets registered to date. Table 2 identifies the control sets registered to date. Re-



member that these assignments, once registered, may never be changed!

Regis. No.	Final Char.	Name
002	4/0	IRV (Intl. Reference Version) Graphics
004	4/1	UK Graphics
006	4/2	US Graphics (ASCII)
008-1	4/3	NATS Main Graphic Set (Finland, Sweden)
008-2	4/4	NATS Additional Set (Finland, Sweden)
009-1	4/5	NATS Main Graphic Set (Denmark, Norway)
009-2	4/6	NATS Additional Set (Denmark, Norway)
010	4/7	Swedish Basic Graphics
011	4/8	Swedish Graphics for Names
013	4/9	JIS Katakana Graphics
014	4/10	JIS Roman Graphics
015	5/9	Italian Graphics
017	5/10	Spanish Graphics
018	5/11	Greek Graphics
019	5/12	Latin-Greek Graphics
021	4/11	German Graphics
025	5/2	French Graphics
027	5/5	Latin-Greek Mixed Graphics (Greek Capitals only)
031	5/8	Greek Alphabet Set for Bibliographic Use

For a G0 set the ESCape sequence is
ESC 2/8 plus the final shown.
For a G1 set the ESCape sequence is
ESC 2/9 plus the final shown.

Table 1. Registered Graphic Character Sets

Regis. No.	Final Char.	Name
001	4/0	ISO 646 Controls
007	4/1	Scandinavian Newspaper Controls
026	4/3	IPTC Controls

The ESCape sequence for a C0 set is
ESC 2/1 plus the final shown.

Table 2. Registered Control Character Sets

The registry set is available from AFNOR for approximately 172 French francs, say $35. It would be vital for an equipment or software manufacturer to have it, and it comes in a beautiful 4-ring binder symbolizing worldwide interchange compatibility. But the summary provided here will fill most needs.

. . .the work I had to do to compact the standard, trying to make it understandable, turned up more than unreadability. So it's back to the drawing board, perhaps for a considerable period of time. . .it's sometimes useful to have symbols whose meaning you can reassign without harm to programming languages. . .

CONTENT OF THE EXTENDED SETS

Figure 1 shows, against the ISO Code, International Reference Version, how the other graphic sets differ in the column/row positions shown. The rows are keyed to Table 1, reminding you that ASCII is "006", or "ISO 646-006".

From this figure we can see that many countries need accented letters as individual characters, not compound via BS (BackSpace). This is particularly true for the double sets 008 and 009, for Scandinavian newspaper transmission, which have characters that cannot be made from ASCII in compound form. For example — Ring-A, a solid, and the angle open and closed quotes.

col	02				03		04	05					06	07			
row	01	02	03	04	10	15	00	11	12	13	14	15	00	11	12	13	14
002	!	"	#	¤	:	?	@	[	\	]	^	_	`	{	|	}	¯
004			£	$													
006				$													~
008-1				$			ua	Ä	Ö	Å	■		ub	ä	ö	å	
009-1		≪	≫	$			ua	Æ	Ø	Å	■		ub	æ	ø	å	
010								Ä	Ö	Å				ä	ö	å	
011							É	Ä	Ö	Å	Ü		é	ä	ö	å	ü
014				$					¥								
015			£	$			§	°	ç	é			ù	à	ò	è	ì
017			£	$			§	¡	Ñ	¿					ñ	ç	~
018			£	$			'										
019			£	$													¨
021				$			§	Ä	Ö	Ü				ä	ö	ü	ß
025			£	$			à	°	ç	§				é	ù	è	¨
027	Ξ		Γ		Ψ	Π	Δ	Ω	Θ	Φ	Λ	Σ					

15a 15b

Figure 15. Registered Graphic Character Substitution

008-2 and 009-2 are shown in Figure 15. Here these are not exceptions from the IRV, but rather the only graphics assigned in the set. The additions are necessary to set type for newspapers throughout Scandinavia. See the Crossbar-D, Crossbar-O, the A-E ligature, and the Icelandic Thorn.

col	04			05				06			07			
row	01	04	05	00	05	11	12	01	04	05	00	05	11	12
008-2	À	Ð	É	Þ	ü	Æ	Ø	à	đ	é	þ	ü	æ	ø
009-2	À	Ð	É	Þ	ü	Ä	ö	à	đ	é	þ	ü	ä	ö

Figure 15. Registered Additional Graphic Sets

15b

Figure 15a doesn't show Set 031 because it deviates more and is not of that much general interest. It doesn't show the Japanese Katakana set because that is completely different from the IRV. In fact, Japanese Industrial Standard C6220-1969 is an 8-bit coded set with the IRV (see Set 014 for the dollar and yen signs) in the lower (bit 8 = 0) portion, and Set 013 in the higher portion, with space reserved for future additional controls. This Set 013 is shown in Figure 2. It is shown in its high-order position, to indicate the card codes at the same time.

Figure 2 also shows the Cyrillic set of the USSR state standard GOST 13052-67, but it is not half of an 8-bit set as the Japanese do it. Rather it is another page of extensions. After SO (Shift Out) is used, the Russian register is operative. Following SI (Shift In) it is the IRV. Although this set has no registry number now, it was submitted recently by ECMA, and we expect an assignment soon. By the way, both Katakana and Cyrillic are shown in their OCR font.

Figure 3 shows the contents of the registered control sets. Set 007 serves as control set for the graphic sets 008-1,2 and 009-1,2, for Scandinavian newspaper transmission. And set 026 is the control set for the worldwide newspaper transmission, defined by the IPTC (International Press Telecommunications Council). The 18 control positions not shown, and those where there is no entry, are the same as in the International Reference Version (646-001).

These newspapers are driving composition equipment, not line printers, so they don't need VT and FF. Their set is already defined, so they don't need SO and SI. They have (properly) assigned meaning to three device controls. And they're probably not doing payroll, so they don't need the four information separators. But they do transmit, and instead of choosing their own functions and placement they have chosen to be a registered variant of the ISO Code. And all variants within this controlled and registered cluster can at least recognize each other, even if they can't print it!

$b_8\,b_7\,b_6\,b_5$ →		1000	1001	1010	1011	1100	1101	1110	1111
$b_4\,b_3\,b_2\,b_1$ ↓	CON TROL	8	9	10	11	12	13	14	15
0000	0				–	ダ ю	Ξ п	Ю	П
0001	1			｡	ｱ	ﾁ а	ﾑ я	А	Я
0010	2			｢	ｲ	ﾂ б	ﾒ р	Б	Р
0011	3			｣	ｳ	ﾃ ц	ﾓ с	Ц	С
0100	4			､	ｴ	ﾄ д	ﾔ т	Д	Т
0101	5			･	ｵ	ﾅ е	ﾕ у	Е	У
0110	6		ｦ	ｶ	ｶﾞ	ﾆ ф	ﾖ ж	Ф	Ж
0111	7		ｧ	ｷ	ｷﾞ	ﾇ г	ﾗ в	Г	В
1000	8		ｨ	ｸ	ｸﾞ	ﾈ х	ﾘ ъ	Х	Ъ
1001	9		ｩ	ｹ	ｹﾞ	ﾉ и	ﾙ ы	И	Ы
1010	10		ｪ	ｺ	ｺﾞ	ﾊ й	ﾚ з	Й	З
1011	11		ｫ	ｻ	ｻﾞ	ﾋ к	ﾛ ш	К	Ш
1100	12		ｬ	ｼ	ｼﾞ	ﾌ л	ﾜ э	Л	Э
1101	13		ｭ	ｽ	ｽﾞ	ﾍ м	ﾝ щ	М	Щ
1110	14		ｮ	ｾ	ｾﾞ	ﾎ н	゛ ч	Н	Ч
1111	15		ｯ	ｿ	ｿﾞ	ﾏ о	゜ —	О	

Figure 2. Katakana and Cyrillic Sets

Position	IRV 001	007	026	
0/09	HT	FO	FO	Format Control
0/11	VT	ECD	ECD	End (a typographical) CommanD
0/12	FF	SCD	SCD	Start (") CommanD
0/13	CR	QL	QL	Quad Left
0/14	SO	UR		Upper Rail
0/15	SI	LR		Lower Rail
1/01	DC1		Font 1	Change to normal
1/02	DC2		Font 2	Change to italic
1/03	DC3		Font 3	Change to bold
1/08	CAN	KW	KW	Kill Word (through previous space)
1/12	FS	SS	SS	SuperShift
1/13	GS	QC	QC	Quad Center
1/14	RS	QR	QR	Quad Right
1/15	US	JY	JY	JustifY

Figure 3. Registered Control Character Substitution

CODE EXTENSION IN ACTION

To illustrate the operation of code extension, let's imagine some equipment that may not exist now:

- A microfiche reader with automatic location controls.
- A microfiche with ASCII (the 8-bit form) on the first two pages, the other pages containing other sets such as Katakana, Cyrillic, Arabic, Greek, Hebrew, mathematical symbols, astronomical symbols, etc. Also, symbol sets for selecting typestyles, weights, rotations, sizes, and elongations.
- A display screen for the microfiche; it is touch-sensitive and generates 7-bit codes according to location touched on the display.
- As an alternative, keyboard tops with fibre optic bundles molded in as a matrix, so that the keytops can be lighted with different symbols as selected.

Now imagine that we are writing an astrology book:

- Type

```
Those of you born under the sign of Aries (
```

- Depress the "astro" key on the special keyboard
- Notice the shift in display for the fiche screen and/or the keytop lighting
- Touch the Aries symbol on the screen (or the keytop)
- Depress SI (Shift In) on the special keyboard
- And return to typing the rest of the sentence

```
) will find this month ...
```

Now imagine what a computer would do to the input stream in driving photocomposition equipment. The "astro" key generated an ESCape sequence for an astronomical graphic symbol set that would have been registered by AFNOR. When the input parser recognizes ESC, it analyzes the following characters, and then calls this set of character formation methods from the backup store, generates the character shape for Aries according to the character code after the final character, notices SI, and returns to normal mode.

Now we can envision how all of the world's printed material can be stored in machine-readable form, and interchanged recognizably!

ALTERNATE CONTROLS

Work has been in progress for several years to develop a companion standard for controls for devices such as CRT terminals. In the US this is contained in the ANSI document BSR X3.64, Additional Controls for Character Imaging. In a similar form, this C1 set is before the Codes Committee of ISO Technical Committee 97 (Computers and Information Processing) as document 2 N 868, for consideration at its 1978 May 24-26 meeting.

I had hoped to give the essence of this work ~~in this installment~~. There were only two negative votes in X3, which one could presume might be answered. Unfortunately, the work I had to do to compact the standard, trying to make it understandable, turned up more than unreadability. It turned up many logical flaws and ambiguities. So it's back to the drawing board, perhaps for a considerable period of time.

Figures 4a through 4e will give, however, some flavor of the controls under consideration.

Figure 4 shows the controls of Format Type (FT) 1 and 2. Format 1 is either the single character of the 8-bit set, shown in the first column as "Ce", or the 2-character sequence of the type "ESC Fe", where Fe is a final character taken from 4/00 to 5/15, and whose column designation is 4 less than Ce. I.e., in an 8-bit code, INDex would be 8/04. In a 7-bit code it would be ESC 4/04. Format 2 is of the type "ESC Fs", where Fs is a final taken from 6/00 to 7/14.

Figures 4b through 4e show controls with formats beginning with the control "CSI", defined in Figure 4a to be either 9/11 (in the 8-bit set) or "ESC [" (in the 7-bit sets). The six possible formats are:

```
3a = CSI Pn F                4a = CSI Pn I F

3b = CSI Pn ; Pn F           4b = CSI Pn ; Pn I F

3c = CSI Ps F                4c = CSI Ps I F
```

Pn stands for numeric parameter(s), Ps for a variable number of selective parameters separated by semicolons. The type 4 formats differ from type 3 only in inserting the intermediate character 2/00 just prior to the final.

In the figures, the parameter value enclosed in parentheses is the default value. That is, if the parameters are not actually inserted, i.e., being null, then the effect is the same as if the default value(s) were inserted.

To give an example of how these controls operate, look in Figure 4d for the second mnemonic, SGR (Select Graphic Rendition). It is represented first by CSI, the Control Sequence Introducer, the parameter, and the final 6/13. This means that when the 4-character string

```
ESC [ 6 m
```

is encountered, it should turn on rapid blink in the field(s) specified on your video screen.

```
AL  = Active Line (containing AP)
AP  = Active Position (where the cursor is)
EF  = Editor Function
FE  = Format Effector
HT  = Horizontal Tabulation
IN  = INtroducer
PAD = Primary Auxiliary Device
RD  = Received Datastream
SAD = Secondary Auxiliary Device
SD  = String Delimiter
VT  = Vertical Tabulation
QA  = Qualified Area (defined by DAQ, SPA, EPA)
rfs = reserved for future standardization
```

Abbreviations for Figures 4a through 4e.

Ce	FT	Type	Param	Mnem	Name
8/00-03	1				(rfs)
8/04	1	FE		IND	INDex
8/05	1	FE		NEL	NExt Line
8/06	1			SSA	Start of Selected Area
8/07	1			ESA	End of Selected Area
8/08	1	FE		HTS	Horizontal Tabulation Set
8/09	1	FE		HTJ	Horiz. Tabul. with Justification
8/10	1	FE		VTS	Vertical Tabulation Set
8/11	1	FE		PLD	Partial Line Down
8/12	1	FE		PLU	Partial Line Up
8/13	1	FE		RI	Reverse Index
8/14	1	IN		SS2	Single Shift 2
8/15	1	IN		SS3	Single Shift 3
9/00	1	SD		DCS	Device Control String
9/01	1			PU1	Private Use 1
9/02	1			PU2	Private Use 2
9/03	1			STS	Set Transmit State
9/04	1			CCH	Cancel CHaracter
9/05	1			MW	Message Waiting
9/06	1			SPA	Start of Protected Area
9/07	1			EPA	End of Protected Area
9/08-10	1				(rfs)
9/11	1	IN		CSI	Control Sequence Introducer
9/12	1	SD		ST	String Terminator
9/13	1	SD		OSC	Operating System Command
9/14	1	SD		PM	Privacy Message
9/15	1	SD		APC	Application Program Command

Fs	FT	Mnem	Name
6/00	2	DMI	Disable Manual Input
6/01	2	INT	INTerrupt
6/02	2	EMI	Enable Manual Interrupt
6/03	2	RIS	Reset to Initial State

18a

Figure 4a. Controls for Character-Imaging Devices

Final	FT	Type	Param	Mnem	Name
4/00	3a	EF	(1)	ICH	Insert CHaracter
4/01	3a	EF	(1)	CUU	CUrsor Up
4/02	3a	EF	(1)	CUD	CUrsor Down
4/03	3a	EF	(1)	CUF	CUrsor Forward
4/04	3a	EF	(1)	CUB	CUrsor Backward
4/05	3a	EF	(1)	CNL	Cursor Next Line
4/06	3a	EF	(1)	CPL	Cursor Preceding Line
4/07	3a	EF	(1)	CHA	Cursor Horizontal Absolute
4/08	3b	EF	(1;1)	CUP	CUrsor Position
4/09	3a	EF	(1)	CHT	Cursor Horizontal Tabulation
4/10	3c	EF		ED	Erase in Display
			(0)		From AP to end (inclusive)
			1		From start to AP (inclusive)
			2		All of display
4/11	3c	EF		EL	Erase in Line
			(0)		From AP to end (inclusive)
			1		From start to AP (inclusive)
			2		All of line
4/12	3a	EF	(1)	IL	Insert Line
4/13	3a	EF	(1)	DL	Delete Line
4/14	3c	EF		EF	Erase in Field
			(0)		From AP to end (inclusive)
			1		From start to AP (inclusive)
			2		All of field
4/15	3c	EF		EA	Erase in Area
			(0)		From AP to end (inclusive)
			1		From start to AP (inclusive)
			2		All of QA
5/00	3a	EF	(1)	DCH	Delete CHaracter
5/01	3c			SEM	Select editing Extent Mode
			(0)		Edit in display
			1		Edit in AL
			2		Edit in field
			3		Edit in QA
5/02	3b		(1;1)	CPR	Cursor Position Report
5/03	3a	EF	(1)	SU	Scroll Up
5/04	3a	EF	(1)	SD	Scroll Down
5/05	3a	EF	(1)	NP	Next Page
5/06	3a	EF	(1)	PP	Preceding Page
5/07	3c	EF		CTC	Cursor Tabulation Control
			(0)		Set HT stop at AP
			1		Set VT stop at AL
			2		Clear HT stop at AP
			3		Clear VT stop at AL
			4		Clear all HT stops in AL
			5		Clear all HT stops in device
			6		Clear all VT stops in device
5/08	3a	EF	(1)	ECH	Erase CHaracter
5/09	3a	EF	(1)	CVT	Cursor Vertical Tabulation
5/10	3a	EF	(1)	CBT	Cursor Backward Tabulation

Figure 4b. Controls for Character-Imaging Devices

Final	FT	Type	Param	Mnem	Name
6/00	3a	FE	(1)	HPA	Horizontal Position Absolute
6/01	3a	FE	(1)	HPR	Horizontal Position Relative
6/02	3a		(1)	REP	REPeat
6/03	3a		(0)	DA	Device Attributes
6/04	3a	FE	(1)	VPA	Vertical Position Absolute
6/05	3a	FE	(1)	VPR	Vertical Position Relative
6/06	3b	FE	(1;1)	HVP	Horiz. and Vertical Position
6/07	3c	FE		TBC	TaBulation Clear
			(0)		Clear HT stop at AP
			1		Clear VT stop at AL
			2		Clear all HT stops in AL
			3		Clear all HT stops
			4		Clear all VT stops
6/08	3c			SM	Set Mode
			1	GATM	Guarded Area Transfer Mode
			2	KAM	Keyboard Action Mode
			3	CRM	Control Representation Mode
			4	IRM	Insertion-Replacement Mode
			5	SRTM	Status Reporting Transfer Mode
			6	ERM	ERasure Mode
			7	VEM	Vertical Editing Mode
			8		(rfs)
			9		(rfs)
			10	HEM	Horizontal Editing Mode
			11	PUM	Positioning Unit Mode
			12	SRM	Send-Receive Mode
			13	FEAM	Format Effector Action Mode
			14	FETM	Format Effector Transfer Mode
			15	MATM	Multiple Area Transfer Mode
			16	TTM	Transfer Termination Mode
			17	SATM	Selected Area Transfer Mode
			18	TSM	Tabulation Stop Mode
			19	EBM	Editing Boundary Mode
			20	LNM	Line feed New line Mode
6/09	3c			MC	Media Copy
			(0)		To PAD
			1		From PAD
			2		To SAD
			3		From SAD
			4		Turn OFF copying RD to PAD
			5		Turn ON copying RD to PAD
			6		Turn OFF copying RD to SAD
			7		Turn ON copying RD to SAD

18c

Figure 4c. Controls for Character-Imaging Devices

Final	FT	Type	Param	Mnem	Name
6/10-11					(rfs)
6/12	3c			RM	Reset Mode
					(same parameters as SM)
6/13	3c	FE		SGR	Select Graphic Rendition
			(0)		Primary rendition
			1		Bold, or increased intensity
			2		Faint, decreased intensity, or secondary color
			3		Italic
			4		Underscore
			5		Slow blink (< 2.5/second)
			6		Rapid blink (> 2.5/second)
			7		Negative (reverse) image
			8		(rfs)
			9		(rfs)
			10		Primary Font
			11-19		1st to 9th alt. font (via FNT)
			20		Fraktur
6/14	3c			DSR	Device Status Report
			(0)		Ready, no malfunctions detected
			1		Busy - retry later
			2		Busy - DSR will notify ready
			3		Malfunction - retry
			4		Malfunction - DSR will notify ready
			5		Please report status (DSR or DSC)
			6		Please report AP via CPR
6/15	3c			DAQ	Define Area Qualification
			(0)		Accept all input
			1		Accept no input (protected); do not transmit (guarded)
			2		Accept graphics
			3		Accept numerics
			4		Accept alphabetics
			5		Right justify in area
			6		Zerofill in area
			7		HT stop at start of area (field)
			8		Accept no input (protected); permit transmit (unguarded)
			9		Spacefill in area

Figure 4d. Controls for Character-Imaging Devices

Final	FT	Type	Param	Mnem	Name
4/00	4a	EF	(1)	SL	Scroll Left
4/01	4a	EF	(1)	SR	Scroll Right
4/02	4b	FE	(100;100)	GSM	Graphic Size Modification
4/03	4a	FE		GSS	Graphic Size Selection
4/04	4b	FE	(0;0)	FNT	FoNT selection
			(0;0)		Primary font
			1;0		First alternative font
			...		...
			9;0		Ninth alternative font
4/05	4a	FE		TSS	Thin Space Specification
4/06	4c	FE		JFY	JustiFY
			(0)		Terminate all justify actions
			1		Fill action ON (text to/from other lines)
			2		Interword spacing
			3		Letter spacing
			4		Hyphenation
			5		Flush left margin
			6		Center text between margins
			7		Flush right margin
			8		Italian form (underscore last)
4/07	4b	FE		SPI	SPacing Increment
4/08	4c	FE		QUAD	Quad
			(0)		Flush left
			1		Flush left, fill with leader
			2		Center
			3		Center, fill with leader
			4		Flush right
			5		Flush right, fill with leader

Figure 4e. Controls for Character-Imaging Devices

Code	Symbol	Code	Symbol
10/00	(same as 02/00)	11/00	Large circle
10/01	Opening double quote	11/01	Dagger
10/02	Closing double quote	11/02	Superior (superscript) 2
10/03	Club suit	11/03	Superior (superscript) 3
10/04	Diamond suit	11/04	Rectangle
10/05	Heart suit	11/05	Parallel
10/06	Spade suit	11/06	Partial derivative
10/07	Closing single quote	11/07	Lower left corner, floor
10/08	Is implied by	11/08	Upper left corner, ceiling
10/09	Implies	11/09	Upper right corner
10/10	Multiply	11/10	Lower right corner
10/11	Plus or minus	11/11	Perpendicular
10/12	Nabla, or del	11/12	Less than or equal
10/13	Em dash	11/13	Not equal, other than
10/14	Radix point	11/14	Greater than or equal
10/15	Divide	11/15	Paragraph mark, pilcrow
12/00	Section mark	13/00	Capital pi
12/01	Double dagger	13/01	Capital psi
12/02	Dot bullet	13/02	Square bullet
12/03	Capital theta	13/03	Capital sigma
12/04	Capital delta	13/04	Integral
12/05	At least one exists	13/05	Capital upsilon
12/06	Capital phi	13/06	Therefore
12/07	Capital gamma	13/07	Capital omega
12/08	Upward arrow	13/08	Downward arrow
12/09	Right arrow	13/09	Left arrow
12/10	Dot product	13/10	Approximately equal
12/11	Degree	13/11	Opening angular bracket
12/12	Capital lambda	13/12	Logical AND
12/13	Register	13/13	Closing angular bracket
12/14	Copyright mark	13/14	Logical NOT
12/15	Capital xi	13/15	Infinity
14/00	Opening single quote	15/00	Small pi
14/01	Small alpha	15/01	Small psi
14/02	Small beta	15/02	Small rho
14/03	Small theta	15/03	Small sigma
14/04	Small delta	15/04	Small tau
14/05	Small epsilon	15/05	Small upsilon
14/06	Small phi	15/06	Check mark, radical mark
14/07	Small gamma	15/07	Small omega
14/08	Small eta	15/08	Small chi
14/09	Small iota	15/09	Logical universal quantifier
14/10	Identically equivalent	15/10	Small zeta
14/11	Small kappa	15/11	Cap intersection
14/12	Small lambda	15/12	Logical OR
14/13	Small mu	15/13	Cup, union
14/14	Small nu	15/14	Overbar
14/15	Small xi	15/15	(same as 7/15)

Table 8. Names of the Additional Graphics, 8-bit Set

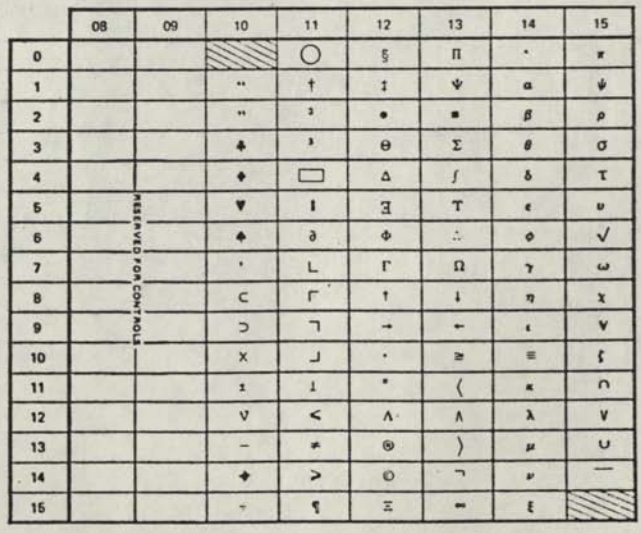


Figure 5. 8-bit ASCII Proposal

CODE EXPANSION

We have seen how ASCII was *extended* by making many related pages of the 7-bit code. It is also possible to *expand* ASCII into an 8-bit code, or even 9-bit and 10-bit if we wished, for that matter. But an 8-bit code is obviously the most logical one to concentrate on, and this has been under development for several years.

The proposed 8-bit Expanded ASCII Code is shown in Figure 5. The identification of the graphic symbols is given in Table 8.

One can observe many interesting things about this set. For example, it has the entire Greek set of small letters except for "omicron", with eleven capitals to go with others from the Roman capitals to complete the Greek set. But apparently the committee didn't follow 646-031, the Greek alphabet mentioned in Table 4. They didn't use the customary ordering "alpha-beta-gamma", the way we learn our "a-b-c's". I suppose it is argued that this set will never be used for language, only math symbols. And 646-027, shown in Figure 12, does not demand the special capital "upsilon" shown in position 13/5. If the Greeks can agree to using a Roman capital "Y" for upsilon, could the Americans?

You'll notice some math symbols, but not enough for APL. In fact, the whole set seems highly slanted to mathematics, rather than business. Of course there are the four corner symbols for forms. Presumably the card suits will strike your eye, and you will wonder why so many other useful symbols were ignored in favor of these. Don't worry, they will always come in handy; it's sometimes useful to have symbols whose meaning you can reassign without harm to programming languages, etc. The committee were obviously bridge players, for spades collate high.

This proposal has not had real public scrutiny yet, and it must be considered no more than a proposal. Presumably X3 will agree about July that it should be sent out for formal public review and letter ballot. My guess is that it will not be adopted in just the form you see here.

FUTURE FOR ASCII

The methods are in place for codifying all symbols that people use. They may be language alphabets, signs, drawing symbols, or controls for equipments. Robots, for example. Satellites are augmenting conventional telecommunications systems, so that one can borrow cheaply and permanently from electronic libraries.

To prepare for this, other sets are being developed for registry, many through ISO Technical Committee 46/1, Automated Documentation. A 2-page mathematical symbol set is near submission, as are African sets. Work is started for Arabic, which will take about 5 sets to handle fully, although there is a commercial subset of 94 graphics. Another C1 set is being proposed for bibliographic controls. It contains four types — annotation controls, filing controls, reference controls, and subject designators. Other C1 sets can come from process control, animation and other graphics applications, etc.

West Germany has proposed a new ISO project on text communication, to harmonize teleconnection of the more than one hundred varieties of typewriters (and keyboards) throughout the world. The extension method of multiple 7-bit codes is ideal for this (8-bit codes imply too many keys or shift combinations for people to use easily.

I am convinced that microcomputer users are going to develop some fantastic applications that will become widespread enough for their special graphic and control sets to be registered. How about a control set or two for sewing machines?

In fact, it is very difficult to think of any general application where one could not find a usage for these registered variants and extensions. □

SHRINKING DATABASES
or
From Entropy to Synergy

R. W. Bemer
Honeywell Information Systems
Phoenix, AZ 85023, US

1978 March 01

Abstract

One can obtain a structured database for relational query by taking a standard framework and entering data into it. But what of the much larger amount of data that exists in mechanized form but is not structured? Will it have to be re-entered at enormous expense?

This paper shows how existing databases can have structure added at modest cost. It is the account of a real conversion of a real, operating database, and the things that happened to it because of that conversion:

- Physical storage space was reduced to one-eighth of the original requirement.
- It became searchable 30 times faster than the original.
- It became subject to relational queries.
- It was made "clean".
- It was suitable as direct publishing input via photocomposition.

Admitting that the original database was 12 megacharacters in size, why is this paper appropriate to a conference on very large databases? Because the identical methods can probably be applied to that class of database with corresponding improvements and reductions. After all, what changed the classification from large to very large? An increase of a factor of 10, perhaps? If so, the process described here would put us back in the "only large" class. And whatever one does to databases, improvements and reductions of this magnitude are always helpful and welcome.

The second admission is that neither IDS nor the DBTG language of CODASYL was used. Instead, the TEX language [1] was used exclusively, 100% in timesharing mode. This language has been submitted to ANSI Standards Committee X3J6, Text Processing. Some of the processes described in this paper have their TEX programs inserted; the reader may ignore them if desired.

The Source Database

The database was one of several used by Underwriters Labs, a product certification organization known worldwide. It consisted of four fields per entry, organized as a linear file: product code, product classification, the complete address of the product manufacturer, and the UL office certifying the product. We'll call it the manufacturer file. Although this paper describes mistakes found in that database, it was nevertheless a far better controlled and cleaner database than today's average.

The processes applied will be described individually. They are summarized here:

1. Fixed fields/records were changed to variable length by use of appropriate dividers and the elimination of multiple blanks.

2. All uppercase letters were converted to lowercase.

3. Selective uppercase reconstitution was applied.

4. The existing minimum (often inconsistent) punctuation was upgraded to be completely and pleasing acceptable for high-quality photocomposition.

5. By appropriate alignment and other manipulation, organizational entries were clustered and analyzed. These were then ordered alphabetically and by state to further refine the clustering. Duplicate successive entries had the manufacturer address field replaced by a double quote character to indicate identity.

6. A new file was created, assigning a token to each unique manufacturing entity.

7. These tokens then replaced all such address entries in the original database.

When received, the database consisted of 24 416 lines, each of 252 characters. The last three lines (being filler) were deleted, leaving 24 413 rated products by many manufacturers. For efficiency, it was broken into five separately-named files by moving the pointer forward 5000 lines, deleting all the remainder, saving as a new file, bringing the original in again, deleting 5000 lines, forward 5000 again, etc. It was then noticed that the starting keys were not easy to remember, so a slight adjustment was made by cut and paste, so that the last four files started with the primary keys E34000, E50000, E60000, and MH9000.

Original occupancy in the HIS Level 66 system was 6717 "little links", or "llinks". It consisted entirely of uppercase characters, suggesting keypunch origination, and was blank-filled to represent card images. Processing on their present equipment is dependent upon this arrangement.

The original database had no direct connection with the creation of the UL product publications, either the books or the card files. That publication was a separate process.

The original manufacturer file also contained a last field of two characters (Ch, Nk, Me, Sc, and Ta), indicating the UL office (city) handling the product and its classification. This interfered with tokenmaking, because a particular plant often made products handled by more than one UL office. The field was deleted, being reconstructable at any time from the original file. Moreover, as a field of constant length, it would be better located in the beginning area of the line, before variable fields occurred.

The Shrunken Database

Original entries (252 characters each) looked like:

```
E12345   QUTY  CARLYLE COMPRESSOR CO DIV OF ... CARRIER
      CORP ... P O BOX 1085 ... SYRACUSE NY13201 ...
```

Compaction, case modification, and punctuation brought them to this form:

```
E12345   QUTY  Carlyle Compressor Co., Div. of\Carrier ...
     Corp.\P O Box 1085\Syracuse, NY 13201|
```

And token replacement yielded this:

```
E12345   QUTY  [CAR2]
```

Entries of the modified database are now 19 characters each, instead of 252. But there is the auxiliary file to decode the tokens, at about 70 characters per entry.

1. Initial Reduction

First all occurrences of 8 blanks were replaced by reverse slash. Then all groups of slash, two spaces were replaced recursively with reverse slash only, to ensure that the (up to 6) blanks in the first 15 positions of each line were preserved. This process brought the storage requirement down to 2079 "llinks" (still in five separately-named files). This process was only a guess at the line separation points, and it turned out to be bad for certain conditions. These were corrected globally as they became apparent. At this point some multispaces remained from the collapsing of the fixed-field format. A general pattern was detectable, but no algorithm. Inspection was required.

Then a constant field of "00" was removed from end of line, replaced by a vertical bar as EOL marker, a very convenient device for future operations.

```
a;*:"#" b rs:"\00#";*:"|" b
```

The operation tally showed that it failed 3 times, which indicated 3 bad data lines in the original file. One was correctable by referencing another line with the same address; two were not.

At this point, storage demand stood at 1775 "llinks", 26% of original. It had taken 4 hours machine time in timesharing, at a total computer cost of $35.

2. Conversion to Lower Case

The first 15 characters in each line had to be uppercase. A TEX program was written to process the other characters selectively, applying known logic. This involved a character-by-character modification of a copy of each line, which then replaced the line. It was well that timing instructions were incorporated deliberately in this program. The cost and elapsed time for this method would be horrendous, perhaps 20 times that of brute force replacement by editor commands, with subsequent recorrection.

Then we found an efficient command library program called "lower". This operation took 10 minutes to convert all capital letters to lowercase, in the 24 413 lines of 85 characters.

3. Recapitalization

Having made all letters lowercase, the problem of initial and other desired caps remained. Some of the processes were:

- All letters following a space, reverse slash, left parenthesis, double quote, or hyphen were changed back to upper case. All lines containing left parentheses were then printed to recheck. For example, "(Formerly I-T-E)" went back to lowercase "f". And "(Usa)" had to go to "(USA)".

- State codes were corrected to caps by a 51-step program "states", one line of which is:

```
rs:" Az8";*:" AZ 8" back
```

- Street directions were corrected by a 4-step program "compass", one line of which is:

```
rs:"Se ";*:"SE " b rs:"Se\";*:"SE\" b
```

- "And", "At", and "Of", if delimited by spaces, were all reduced again to lower case.

- Names starting "Mc" or "Mac" (except the children's classic (MacHinery") had the next letter capitalized. The program "mc" required one program step for each such letter.

- Wellknown initials such as "Trw" and "Itt" were changed to all caps if followed by a space.

Special problems occurred for names and addresses that were not English:

- Certain words delimited by blanks should not be capitalized -- a, bis, de, des, du, et, les, pour, rue, sur, und, fur (but gets a hyphen) ...

- If the second character is an apostrophe, the first is lowercase, and the letter following the apostrophe is upper case.

- If the first character is "s", and the second a slash, the initial "s" is kept lowercase, and the letter following the slash is capitalized. Example: Neuilly s/Seine.

- " s a" or " sa" all go to " SA", for Societe Anonyme; " Spa" and " S P A" to " SPA"; " Gmbh" to " GMBH".

- And of course " Tv" goes to " TV".

4. Repunctuation

Consistent punctuation was one prerequisite to structuring the database. If the name of the same company was "Jones Inc." in one entry and "Jones, Inc." in another, the automatic matching would not occur via a sort. Of course, it would have been nice if other automatic identity detection methods were available. But that would be a very complicated program, and we would be in the fuzzy set business. For example, I ran a concordance on a US Government database, and found that three high-level executives of one company had remarkably similar names -- Wohlegemuth, Wohlgemuth, and Wolgemuth!

The program tells the story most simply. ";*:" means "replace all occurrences of the previous quoted string with the following quoted string":

```
rs:"Co\";*:"Co.\" b rs:" Co ";*:" Co. " b rs:"Co,";*:"Co.,"
b rs:"Rd ";*:"Road " b rs:"Rd\";*:"Road\" b
rs:"Rt ";*:"Route " b rs:"Rte ";*:"Route " b
rs:"Ln ";*:"Lane " b rs:"Ln\";*:"Lane\" b
rs:"Dr ";*:"Dr. " b rs:"Dr\";*:"Dr.\" b
rs:"St\";*:"St.\" b rs:"St ";*:"St. " b
rs:"Corp\";*:"Corp.\" b rs:"Corp ";*:"Corp. " b
rs:"Inc\";*:"Inc.\" b rs:"Inc,";*:"Inc.," b
rs:"Ltd\";*:"Ltd.\" b rs:"Ltd ";*:"Ltd. " b
rs:"Div ";*:"Div. " b rs:"Div\";*:"Div.\" b
rs:"Div. Of";*:"Div. of" b
rs:"Co. Inc";*:"Co., Inc" :"Co. Ltd";*:"Co., Ltd" b
rs:"Hwy ";*:"Hwy. " b rs:"Hiway";*:"Hwy." b
rs:"Dept ";*:"Dept. " b rs:"Dept\";*:"Dept.\" b
rs:"Terr ";*:"Terrace " b rs:"Terr\";*:"Terrace\" b
rs:"Ave ";*:"Ave. " b rs:"Ave\";*:"Ave.\" b
rs:"Mfg ";*:"Mfg. " b rs:"Mfg\";*:"Mfg.\" b rs:"Mfg,";*:"Mfg.,"
b rs:"Blvd ";*:"Blvd. " b rs:"Blvd\";*:"Blvd.\" b
rs:"Ct ";*:"Court " b rs:"Ct\";*:"Court\" b
rs:"St ";*:"St. " b rs:"Ft ";*:"Ft. " b rs:"Mt ";*:"Mt. " b
rs:"Dba";*:"dba" b rs:"L I,";*:"LI," b
```

File cleanup and modification to this point took about $800 of machine time, or about 3 cents per entry, which is still cheaper than rekeying and revalidating. But that was because we were feeling our way in a new area. With what is now known, it could be redone for $100.

5. Clustering and Analysis

The work was started with companies that had a large number of entries, as determined by visual inspection of a listing. Their entries were cut from the file on the basis of a single string clue. This worked well enough, but as the ore became lower grade, as it were, we mined for five strings simultaneously in one pass through the file. This was a mistake. It was ineffective. Not enough tokens were inserted to shrink the file for further search. What really worked was an oldfashioned sort of the entire database.

Entries of multicompany corporations were put into separate files, one for each corporate entity. Each of these files were processed by a program "states", which made successive extraction in alpha order on the state. This was a most important and successful procedure. The clustering enabled detection of such cases as:

- Mistakes in address components -- PO Boxes, ZIPcodes, spellings for streets, company components, cities, etc.
- Inconsistencies in components, e.g., U S Rte 60 in one entry, just Rt 60 in another.
- Omissions of organizational detail in one entry, compared to another at the identical address.

Most of the name and address fields started with the company name, all in the same column so they could be ordered simply. But over 1500 of the 24500 entries were "A, Div. of B", or a subsidiary, etc. These were separated from the main file(s) by:

```
cuts:"of\";*
```

The sort had to be provided with a fixed field format, to align "B" for ordering. The program was:

```
s="                " space=s,s,s,s,s
!loop if *eof out:"Done." return
scan:*cl:"\" a=(space,*l)[70
L=a,*m,*r r:*cl:L f;1 goto !loop
```

Then the sort was called:

```
sort *;*:(A71,A30)(A2)
```

"A71" means that the first alpha field is 71 characters, our controlled width of 70 characters from the previous program, plus the reverse slash separator. "A30" specifies the second alpha field, for "B". "A2" means order (ascending) on the second alphanumeric field. After ordering, the file was returned to original condition by:

```
!loop if *eof out:"Done." return
L=*cl>' " " r:*cl:L f;1 goto !loop
```

This searched each line for the first non-space and saved the righthand part.

A check of the remainder file, supposedly individual manufacturers, showed deficiencies in the indication of company structure. Some entries for Carlyle Compressor Co., -- same address -- did not show it as a division of Carrier! This occurred for perhaps twenty companies. It was more difficult to group parent companies with their several divisions, subsidiaries, etc.

Now the file was ready to detect and replace duplicates by ditto marks:

```
!undupl split:*cl:15 old=*r,/!/ L=*l,old r:*cl:L f;1
!loop if *eof out:\Done.\ return
split:*cl:15 new=*r,/!/
if new:eqs:old L=*l,/ "/ r:*cl:L old=new f;1 goto !loop
L=*l,new r:*cl:L old=new f;1 goto !loop
```

Manufacturers with a single plant were then cut from the remainder of the file as ordered on name-address. Tokening was repeated, except that this time the digit tag was a constant "1" (anticipating growth), and a new token was requested each time a new entry without ditto marks appeared. Again the replacement files were ordered, broken, adjoined, and reordered into position for replacement.

This was done for clusters of four or more products from the same plant. Now we had to ask whether groups of two and three should be permitted to be lost. If so, the consistency advantage might disappear. And so would some relationships.

It was decided that the added storage demand and processing time was not enough to offset the advantages of tokening everywhere possible. But this increased greatly the number of tokens required (and duplicates cannot be allowed), and we did not wish to go to 4-letter tokens. The decision was to play license plate. Whereas:

```
[CAR6]
```

indicated the sixth plant of Carrier Corp.,

```
[2XYZ]
```

indicated that XYZ could be decoded from a secondary token list, and that there were two products for that plant. This would have to be changed to [3XYZ] to add another product, but that can be done automatically. So can the transfer to the primary token list if it went to four products.

6. Making Tokens

The program "maktoken" asked for a proposed token of two or three capital letters. The US Stock Exchange representations were used when possible. The existing token equivalence file was searched to see if that token was already used. If so, a new token was requested. This program:

- Copied from the file all entries with full addresses (no ditto marks). Stepping the tag count by 1, it made line entries from these to add to the token equivalence file.

```
CAR2 = Carlyle Compressor Corp., Div. of\Carrier ...
CAR3 = Carrier Airconditioning\ ... , etc.
```

- Modified all entries so that either address or ditto mark was replaced with the token:

```
E12345   QUTY  [CAR2]
```

7. Replacement by Tokens

These tokened entries were accumulated in a replacement file. When enough were accumulated (500 to 1000 in this case), the replacement file had the first space in each entry replaced by the Unit Separator character (US) of ASCII. Then it was split into five parts to correspond to the main parts of the database. The five file pairs, replacement and database, were adjoined and ordered on the first field (product code). Because US collates lower than space, the ordering stuck the replacement lines just before the lines that they were to modify (like sticking changes by hand into a box of cards). This program transferred the token to the real entry:

```
!work b dup="n"
!loop fs:*us if *eof out:"Done." return
scan:*cl:*us token="[",*r,"]" d
f:*l if *eof b;1 line=*l,"*",*r i:*cl:line p b goto !loop
!dupl
if *lcl:lt:25 f;1 f:*l dup="y" out:"Dupl on ",*l goto !dupl
split:*cl:15 line=*l,token r:*cl:line out:line
if dup:eqs:"y" goto !work
goto !loop
```

ZIPcodes as Tokens

When taken to all digits of detail, the US ZIPcode defines a location almost unambiguously. There may be many ZIPcodes for a single city, which is no problem, but sometimes more than one named place has the same ZIPcode. There is also the useful property that the first digit is the same for all locations in a given state, with a general progression from "0" in the East to "9" in the West.

The full form, such as "Phoenix, AZ 85023", may be useful or mandatory in a working address, but it is not usually necessary to keep anything but ZIPcode in a database, once the entry is validated via redundancy. A separate ZIPcode directory is needed to make the expansion when human output is required. Thus ZIPcode will serve as an additional token type.

The UL database was measured to see how much storage demand could be reduced by tokening the cities by ZIPcode. This was the process:

- This program kept only city-state-zip for US addresses, inverting to zip-city-state to yield a directory of codes actually used in the database:

```
!loo if *eof out:"Done" return
scanr:*cl:"\" a=*r<*n if a:gt:2 d;1 goto !loo
splitr:(*r'[1):5 L=*r,(*l'<" ") r:*cl:L f;1 goto !loo
```

- 2121 of non-US were deleted, leaving 22285. This file was saved as a new file, to find a storage demand of 432 "llinks". The average length was 18 characters, which would have been 19 in the original form (space between state code and zip), for an average of 10 for city name. Then a new file was created of ZIPcodes only.

```
!loop if *eof out:"Done" return
splitr:*cl:5 r:*cl:*r f;1 goto !loop
```

- These codes were ordered and duplicates eliminated:

```
sort *;*:(A5)(A1)
call texlib/elimdup
```

501

- 4790 unique ZIPcodes were found, a 4.6 to 1 reduction. Storage demand was 46 "llinks". After more punctuation cleanup, the same process was applied to the ZIPcode directory with city attached:

```
sort *;*:(A22)(A1)
call texlib/elimdup
```

471

- Here the number of unique lines was found to be 5152, but the average went up to 19 characters. Storage demand was 96 llinks. As only 4792 codes were extracted, obviously something was not unique in the city-to-ZIPcode correspondence. These differences were displayed by:

```
al=*cl a=*cl']5 f;1
!f if *eof out:"Task completed." return
b=*cl']5 if a:eqs:b out:al out:*cl,*lf f;1 goto !f
a=b al=*cl f;1 goto !f
```

- Duplicates were analyzed visually out of curiosity:

  One was a dispute about state. Was Lewiston in Maine (ME) or Massachusetts (MA)?

  68 were the result of misspellings, with the misspelled city counting only once, as it was not determined how many times it was actually misspelled in the database.

  231 were actually ambiguities in the ZIPcode system, where cities of different name had the same ZIPcode. Of course it's anyone's prerogative to be addressed at St. Paul, MN, if it is thought that someone else might not find him at Circle Pines, MN -- same ZIPcode. Or Dallas, TX rather than Farmers Branch.

  11 were obvious ZIPcode-city mismatch mistakes. There may have been many more.

  17 resulted from using LI for Long Island in some addresses, and not in others.

  28 were variations in usage (Ft.-Fort) (Mt.-Mount) (Pk.-Park) (Hts.-Heights) (St.-Saint), etc.

  4 were from using both a single character and full name for directions (N-North).

Conclusions on ZIPcodes as Tokens

To use ZIPcode as a token, without ambiguity, requires that the database token be subscripted with an added character, to differentiate among cities with the same code. This will increase storage demand (for the ZIPcode portion only) by perhaps 5%, not 20%, as not all codes will have to go to 6 digits (remember that the database method described here does not require fixed field lengths!).

The storage-saving potential is good, but the more important result is generation of correctly-spelled cities, even though they may be entered erroneously into the database. The full ZIPcode directory can be called upon at entry time, with most corrections being done automatically, perhaps without notice to the enterer. If a mistake cannot be resolved, a prompting query can be issued.

Consistency of abbreviation may also be obtained, and one can choose, for example, whether or not "Long Island" is to be included in those New York State addresses.

In the case of this particular database, already compacted greatly by tokens for the total manufacturer entity, the storage saving is not so great, but the cleanliness is still most attractive. Hardly any of the mistakes detected by ZIPcode examination were caught in the normal tokening process.

Discoveries

Some significant lessons were learned in this work:

- Publishing requirements on a database are met easily by a text processing language, but not by conventional or DBTG methods.

- There are few databases constructed originally to the DBTG specs, compared to those already entered and in existence, demanding conversion and structuring. Moreover, existing databases are almost always "dirty", so that the elements of structure are detectable only via text processing. (Mayford Roark, Director of Systems for the Ford Motor Company, has much database experience [2]. After looking at these results, he said that he had never seen this much structure in a database before).

- There are substantial advantages to having a database as a linear file: clustering for visual inspection, restructuring at will for unforeseen purposes, ability to apply global text editing, and use of standard sorting methods.

- The token method is useful for altering or adding new entries to the database. Directories, printed periodically, can be used by all workers as a shorthand. The compaction allows the database structure to be more visible.

REFERENCES

1. The TEX Subsystem of the Timesharing System, Series 60 Level 66, Honeywell Information Systems, 200 Smith Street, Waltham, MA 02154, Order DF72.
2. Mayford Roark, "Evolution in Computer Systems", Keynote Address, Database Directions II, 1977 Nov (to be published by the US National Bureau of Standards, Gaithersburg, MD 20760.

3.

Acknowledgements

I thank Underwriter Laboratories for the opportunity of working with this database, and the learning that it provided. In particular, Ted Conterio, in charge of their data processing operation.

1977

PRELIMINARY PROGRAM

SEVENTH MEASUREMENT SCIENCE CONFERENCE

FRIDAY, December 2, 1977

Registration – 8:00 to 8:45 AM

Business Meeting – 8:45 to 9:15 AM
Welcome - Dr. Stephen Horn, President, California State University Long Beach, Long Beach, California

Keynote Address – 9:15 to 10:00 AM
Robert Bemer, Honeywell Information System Inc., Phoenix, Arizona

Break – Continental Breakfast and Exhibit Viewing - 10:00 to 10:45 AM

Session 77-1 - 10:45 AM to 12:15 PM

1A NBS Part I - "Overview of DC/Low Frequency Measurements & Standards"

Developer - Dr. Barry Taylor, Chief, Electricity Division, NBS, Washington, D.C.

After the introduction to the Electricity Division of the National Bureau of Standards - its goals, resources and recent accomplishments - the talk will focus on its current research work in the area of DC - LF measurements and standards.

1B "Automated Test Systems"

Developer - Richard Oderwald, HQ USMC, Washington, D.C.

Paper - "Automatic Calibration Revisited"

Frank Capell, Product Planning Manager, John Fluke Manufacturing Co., Mount Lake Terrace, Washington 98043

Two hundred years of experience have been accumulated with automated calibration systems manufactured by the authors company. This experience suggests that much larger use of automation in calibration laboratories can be expected in the future, but is not a panacea. The very important aspects of automated calibration and the subtle tradeoffs between man and machine are described. Implications of the IEEE 488 Standard are discussed.

Paper: "Introduction to Atlas"
Roy Nessen, Hughes Aircraft Co., Culver City, CA 90230

ATLAS is a high level language used in preparation ATE compatible test procedures. A film which describes the characteristics of ATLAS will be shown.

1B Paper: "Automated Test System for Amplifiers"
Robert D. Weir, Boeing Aerospace Co., Seattle, Washington 98124

This paper describes development of a tape programmed automatic test system which checks the performance characteristics of amplifiers. The discussion covers translation of amplifier test requirements to selection of programmable test equipment with adequate stimulus/measurement capability and software requirements for proper test sequencing and test data recording.

Paper: "Microprocessor Device Production Testing"
Edward J. Vignone, Microelectronics Devices Division, Rockwell International, Anaheim, California 92803

Martin H. Gruenthal, Collins Commercial Telecommunication Division, Newport Beach, California 92663

This paper traces the development of production test devices from the basic test requirement through the manufacturing process of testing the device. The discussion covers the translation of product requirements into an Automatic Test compatible language along with a description of the measurement equipment required to implement these tests. Operator interface and production test data requirements are highlighted to round out the discussion.

Paper: "A Second Industrial Revolution is Automated Low Labor Costs, Maintenance (Standards, Calibration, Repair)"
Loeb Julie, Julie Research Laboratories, New York, N.Y. 10023

The technical advances of the Industrial Revolution brought about a ten-fold reduction of the labor costs in manufacturing. As a result of this labor cost reduction, the whole range and variety of modern manufactured products was made available and all of these products can be obtained at extremely low initial purchase prices relative to their complexity. Unfortunately, in 95% of present day plants, the real cost of ownership of the entire plant inventory of modern instrument products is two to three times higher than procurement costs because of traditional, high labor cost maintenance. In view of the enormous ownership costs of the large present day plant instrument inventory, what is needed is a second Industrial Revolution capable of bringing about a ten-fold reduction in the high cost of maintenance.

This paper examines the cost and savings impact of the successful automated low labor cost maintenance (standards, calibration, repair) on the cost of ownership of the fifteen billion dollar instrument inventory (estimated) of present day U. S. Companies and presents data showing how, eventually, every company can benefit from the two revolutions of automated maintenance.

Robert Berner
HIS

ABSTRACTS OF CONTRIBUTED PAPERS

OF THE

FIFTH ANNUAL MOUNTAIN REGION COMPUTING CONFERENCE

WHITE WINROCK MOTEL

ALBUQUERQUE, NEW MEXICO

INTERPRETERS AND THE SOFTWARE LIFE CYCLE

by

Robert W. Bemer
Honeywell Information Systems
Phoenix, Arizona

Every piece of software has four major tradeoff variables - people time and resources, computer time and resources, when it has to be done, and how often the task must be done. The figures of Boehm show the average allocations of people time to be 12% design, 6% coding, 12% testing, and 70% maintenance (e.g., change, correct, expand, add features, etc.)

This good start is short of the total solution, as it considers but one of the four variables, but the term "life cycle" is correct. This paper will study the economics of software production by comparing one vector - a spectrum of applications - against a method consisting of:

- Total use of an interpreter (TEX, in this case).
- Using the interpreter to scope and check the method, then recoding in a standard compiler language for production.
- Using the standard compiler languages exclusively.

Über den Computer in unserer Gesellschaft

Computers and our Society*

Elektron. Rechenanl. 19 (1977), H. 4, S. 167 – 172
Manuskripteingang: 22. September 1976

von R. W. BEMER
Honeywell Information Systems,
Phoenix, AZ, US

Computeranwendungen werden in drei Kategorien eingeteilt, (1) beratend, (2) zur Entscheidung durch den Menschen führend, oder (3) mit Entscheidungen, die ein vorprogrammierter Computer trifft, wenn nicht rechtzeitig eingegriffen wird. Einige Beispiele für Schwierigkeiten selbst in den ersten beiden Kategorien lassen äußerste Vorsicht in der dritten Kategorie geboten erscheinen. Die aus der Raumfahrt gewonnenen Erkenntnisse der Computer-Technologie sind auf den Großteil der Computer-Anwendungen noch nicht übertragen worden. Sowohl gesetzliche wie auch freiwillige (fachliche) Maßnahmen gegen den Mißbrauch werden diskutiert.

Computer usage is classified as either (1) advisory, (2) leading to decisions by humans, or (3) with decisions being taken by a preprogrammed computer unless countermanded in time. Some examples of difficulties even in the first two categories imply that caution in the third is imperative. The computer technology learned from the space effort is not yet transferred to the bulk of computer usage. Both legal and voluntary (professional) measures against misuse are discussed.

Einleitung

Im Jahre 1950, nach meiner Spätschicht bei der Rand Corporation, war ich noch um 8.30 Uhr an einer 604-Konsole damit beschäftigt, eine achtstellige Quadratwurzel aus einer achtstelligen Zahl zu ziehen (zu dieser Zeit war diese Operation bei diesem Gerät noch nicht automatisch). Ein kleiner, rundlicher Mann kam zu mir und fragte, was ich da mache. Ich sagte es ihm. Er fragte mich dann über den Rechner aus, und jeder meiner Antworten folgte eine schwierigere und eindringlichere Frage, bis ich schließlich Mühe hatte, adäquate Antworten zu geben. Er stellte sich nicht vor, aber ich hörte dann später, daß es *John von Neumann* gewesen war.

Dieser Vorfall ist mir natürlich lebendig in Erinnerung geblieben. Ich weiß heute noch, daß er beim Abschied *nicht* sagte: „Verwenden Sie dieses Werkzeug gut zum sozialen Wohl der Menschheit", oder irgend so etwas. Es gab zu jener Zeit in der Computerbranche oder in der Industrie sehr wenige, die soziale Überlegungen anstellten. *Edmund Berkeley* war und bleibt eine Ausnahme. Für die meisten von uns war das bloß eine Zeit, unseren Geist mit phantastischer Geschwindigkeit für bisher nicht dagewesene Fähigkeiten freizumachen. Wir waren verzaubert und begeistert zugleich: der Reiz des Neuen und die ungeheuren Möglichkeiten lockten uns, und so viel wartete darauf, getan zu werden. Da war wenig Zeit für Spekulationen über die möglichen Auswirkungen des Computers auf unsere Gesellschaft oder über das Ausmaß und den Umfang zukünftiger Anwendungen.

Diese Unempfindlichkeit mag auch darauf zurückzuführen sein, daß damals fast ausschließlich Zahlen verarbeitet wurden. Selbst als ich mich im Jahre 1949, zehn Jahre nachdem der erste programmgesteuerte Rechner gebaut worden war, mit diesem Gebiet zu beschäftigen begann, befaßten sich nur einige wenige mit der Zeichenverarbeitung, und es brauchte bis zum Jahre 1956, ehe die Zeichenverarbeitung als geziemende Computerfunktion angesehen wurde.

Ich möchte damit zeigen, daß sich die durch den Computer ermöglichten Anwendungstypen beträchtlich geändert haben; eine Änderung, für die wir miserabel vorbereitet sind. Jedes Werkzeug, das Kraft und Verstärkung vermittelt, kann mißbräuchlich verwendet werden. Ich werde einige Fallgeschichten erzählen, um einige Formen des Mißbrauchs aufzuzeigen und zu erklären, warum sie auch in der nächsten Zukunft nicht aufhören werden. Dann will ich einige Maßnahmen skizzieren, wie der Trend umgekehrt und ein Großteil der mißbräuchlichen Verwendung unterbunden werden kann.

Eine Einteilung der Computeranwendungen

Für die Zwecke dieses Beitrages möchte ich hier eine einfache und vielleicht neuartige Einteilung der Computeranwendung vorschlagen:

1. Anwendungen, die nicht zu Entscheidungen führen, welche den Menschen direkt betreffen.

   Beispiele dafür kommen hauptsächlich aus dem Gebiet des numerischen Rechnens, der frühesten Anwendungsart. Rechenergebnisse, die eine Theorie beweisen oder initiieren könnten; Berechnungen für den Entwurf von Raumfahr-

* Dies ist eine gekürzte Version eines Vortrages bei der NordDataTagung in Kopenhagen vom 15.–17. August 1973, erschienen im Honeywell Computer Journal 8 (1974) no. 1; pp. 49–54. Übersetzt von *R. Reithofer*, Wien. Als mein alter Freund *Bob Bemer* die Erlaubnis für die Übersetzung und deren Erscheinen in den eR gab (er ist auch Herausgeber der Honeywell Computerzeitschrift 1972, Band 6), versprach ich ihm eine gute, aber nicht wörtliche Übersetzung. Wo die gute Absicht schiefging, trifft es mich.

H. Zemanek

zeugen oder Raketen (diese müssen nicht gebaut oder gestartet werden); Spielprogramme oder Programme, welche Gewinn und Verlust mit Strategien verknüpfen, und so fort. Derartige Anwendungen können wir als *beratend* bezeichnen.

2. Anwendungen mit Rechenergebnissen, die zu Entscheidungen durch den Menschen führen.

   Einige dieser Anwendungen können der Einmischung in menschliche Angelegenheiten sehr nahe kommen. So kann zum Beispiel jemandem ein Kredit oder eine Anstellung verweigert werden. Es hat sich in langer Praxis gezeigt, daß die zu treffende menschliche Entscheidung oberflächlich oder gedankenlos sein kann. Es gibt jedoch Abhilfe, wie zeitraubend und schwierig sie auch sein mag, und ganz gleich, welche gesetzlichen Maßnahmen in Gang gesetzt werden müssen, um Menschen vor solchen Situationen zu schützen.

3. Anwendungen, wo der Computer vorprogrammiert wurde, eine Entscheidung und Handlung auszuführen, und diese Handlung auch ausführen wird, wenn nicht rechtzeitig eingegriffen wird.

   Beispiele dafür sind die Überwachung von Patienten durch den Computer, die Steuerung von Atomkraftwerken, Flugüberwachungs- und Flugsicherungssysteme, automatische Transportsysteme oder Automobilbrems- und Antischleudersysteme.

Die Hardware-Entwicklung der letzten drei Jahre, die zu Mikroprozessoren auf Chips führte, verspricht eine gewaltige Steigerung der dritten Anwendungskategorie. Aus diesem Grund müssen wir auf die Richtigkeit und Systemaspekte solcher Anwendungen besonderes Augenmerk legen. Bei Digital-Armbanduhren scheint es weniger kritisch zu sein, bei Autos hingegen können solche Anwendungen extrem kritisch werden. Man denkt dabei sofort an die Servolenkung: eine Wohltat, wenn sie funktioniert, aber eine ausgesprochene Gefahr, wenn die Energieversorgung ausfällt oder abgeschaltet wird.

Das folgende Fragenpaar zeigt das mögliche Dilemma auf:

Gibt es eine Technologie, um Computerkomponenten sehr eng in den menschlichen Bereich zu integrieren?

Ja. Ein Beispiel sind die 1974er Modelle der amerikanischen Autos, die nur dann in Bewegung gesetzt werden können, wenn der Fahrer angeschnallt ist.

Gibt es für solch eine Ebene der Technologie Handbücher für Systementwurf und eine gute Praxis, beziehungsweise bieten unsere Schulen adäquate Einführung und Ausbildung dafür an?

Nachdrücklichst: nein! Und das ist erschreckend genug, für solche Entwicklungen ein Stillhalteabkommen vorzuschlagen, bis wir das Werkzeug besser verstehen.

Nehmen wir zum Beispiel die Meldung über einen bestimmten Vorgang, bevor Sie Ihr Auto starten können. Die Absicht besteht darin, betrunkene Fahrer daran zu hindern, ihr Fahrzeug in Gang zu setzen. Aber stellen Sie sich vor, Sie sind extrem verzweifelt, weil Ihre Frau eben getötet wurde und Sie Ihr verletztes Kind ins Krankenhaus bringen müssen. Könnten Sie in einem solchen Zustand das Auto starten?

Gewiß hat die amerikanische Raumfahrt reichliche Erfahrung darin gewonnen, Computer entscheiden zu lassen, wenn sie dazu fähig sind, und sich vernünftig über sie hinwegzusetzen, wenn sich herausstellt, daß sie falsch programmiert sind oder daß gewisse Funktionsstörungen nicht bedacht worden waren. Wir sehen viele Nachwirkungen der Raumfahrt-Technologie bei Produkten (Hardware) aller Art, aber sehr wenige in der Methodologie – gerade sie aber wäre für die Computeranwendung so eminent wichtig.

Ein Wort der Warnung

Ich möchte in aller Klarheit vorausschicken, daß ich Computer gern habe. Ich glaube, daß sie der Gesellschaft gegenwärtig mehr nützen als schaden und daß dieses Verhältnis noch verbessert werden kann, wenn wir sie vernünftig einsetzen und ihre Verwendung möglichst gut und sinnvoll planen. Wäre ich schicksalsgläubig, wenn ich sagen würde, daß sie gerade rechtzeitig gekommen sind, um uns vor unseren Feinden zu retten.

Eine Hauptschwierigkeit bei der Beurteilung des Wertes eines Werkzeugs ist die Unmöglichkeit, die Verwendung eines Werkzeugs als *absolut* gut oder schlecht zu kategorisieren. Das soll nicht philosophisch sein, sondern uns nur daran erinnern, daß wir diese Urteile von gut und schlecht innerhalb des engen Rahmens unserer Moralbegriffe sehen, welche wiederum durch unser angesammeltes Wissen und durch unsere subjektive Betrachtung des Weltablaufs bedingt sind. Wir haben den Weltlauf in letzter Zeit etwas besser und objektiver verstehen gelernt, nicht weil wir dieses Wissen so sehr suchten, sondern eher, weil uns schmerzlich bewußt wurde, daß alle Elemente unserer Welt mehr oder minder verkoppelt sind.

Ich werde mich daher mit einigen schlechten Computeranwendungen befassen, nur um die Probleme aufzuzeigen, die durch Gesetzgebung, Erziehung und Fachverantwortung gelöst werden müssen, damit uns der Computer besser dienen kann.

Die Macht des Computers

Eine IBM-Anzeige betont: *Betrachten Sie den Computer als Energie!* Theoretisch ist der Computer eine ungeheure Macht im Dienste des Menschen. Aber lassen wir uns nicht durch die Werbung dazu verführen, die Energie des Computers wie elektrische Energie zu betrachten. Computermacht ist Arbeitsmacht, aber sie ist auch Wissensmacht, eine Macht also, wie sie durch die ganze Menschheitsgeschichte sowohl zur Machtausweitung als auch zum Wohle des Menschen verwendet wurde. In einer Zeit, in der die Technologie in die Enge getrieben ist, wird es gut sein, die Gefahren des Computermißbrauchs durch die Schaffung von Vorurteilen in der Bevölkerung gegenüber einem wertvollen Werkzeug in Betracht zu ziehen sowie die Gefahren des Mißbrauchs durch korrupte oder unwissende Beamte.

Es sind keinerlei Fälle bekannt, in denen Computer aus eigenem Antrieb die normale Arbeit stoppen, um dann ohne menschliche Anweisung illegale Handlungen zu begehen. Computermacht ist allgemein verfügbar, und wir dürfen nicht überrascht sein, daß sie von einigen Leuten mißbraucht wird.

Betrachten wir den Fall von *Jerry Schneider* – man kann unbedenklich darüber reden. Er beschreibt, wie er ein elektronisches Auftragssystem anzapfte und Telefongeräte im Wert von ungefähr einer Million Dollar stahl, indem er veranlaßte, daß die Geräte zu einem Lastwagen einer Telefongesellschaft geliefert werden, den er auf einer Auktion gekauft hatte. Das Computerprogramm, das nicht wußte, an wen die Rechnung auszustellen war, ignorierte diesen Posten, weil er unterhalb der Verlustgrenze lag. Nachdem ihn ein Angestellter überführte, verbrachte *Schneider* zwei Monate lang im Gefängnis. Er kehrte dann in sein Fachgebiet zurück: als Berater für Computersicherheit!

Dr. *Henry Bruck* vom Massachusetts Institute of Technology sprach darüber auf der ACM-Tagung 1970 in seinem Vortrag *Die Wiederherstellung des Gleichgewichts* [2]. Seine These war, daß Computer auf Grund ihrer Kosten und Ausbildungsinvestitionen eher zu Werkzeugen der Regierung und großer Unternehmen als zu denen der Allgemeinheit werden. Gegen das Argument, daß Minicomputer, Mikrocomputer und Taschenrechner der Allgemeinheit zu geringen Kosten zugänglich sind, sagte er, die Annahme, daß aus diesem Grunde die Macht des Computers der Allgemeinheit zur Verfügung stünde, sei ein Trugschluß. Schaufeln um drei Pfennig sind nutzlos, wenn man nicht weiß, wie man gräbt, und zudem keine Arme hat. Das Wichtigste ist die Fertigkeit im Gebrauch eines Werkzeugs.

Er war der Meinung, daß es eine unnötige Überspezialisierung wäre, die Ausbildung so abzuändern, daß das Erlernen grundlegender EDV-Kenntnisse (und Problemlösungsmethoden) ebenso starkes Gewicht bekäme wie das Erlernen der eigenen Muttersprache.

Ich bin derselben Meinung. Es gibt viele Möglichkeiten, Computerdienstleistungen durch Stadtverwaltungen und private Unternehmen zu vermitteln. Man kann sich Datenbanken vorstellen, die als Ratgeber dem Menschen dienen. Dem Konsumenten könnte auf vielerlei Art geholfen werden – Produktsicherheit und Wirtschaftlichkeit, Vergleich der Einkaufsmöglichkeiten, Finanzhilfe bei größeren Einkäufen, Erinnerungshilfe für Steuervorauszahlungen usw.

Es gibt also viele Möglichkeiten, ohne daß der Betroffene programmieren lernen muß. Es ist jedoch ein gewisses Maß an *Computer-Bildung („computer literacy")* erforderlich, will man sich bei einer derartigen Verwendung wohl fühlen.

Ein Blick auf die Sünden, die der „Autorität" des Computers entspringen

Wie kein anderes Werkzeug vor ihm (nicht einmal das Auto) wurde der Computer in kürzester Zeit zu einem Gemeinplatz. Das hat zu Verwirrung und Unbehagen geführt, und auch als sie es merkten, konnten die Fachleute nicht viel dagegen tun. Die meisten größeren Werkzeuge hatten, als sie eingeführt wurden, zunächst ihre Hüter und dann ihre Gilden und Innungen, die auf Grund fortschreitender Erfahrung in Gesetz und Praxis jene Sicherheitsklauseln für deren Verwendung einbauten, welche auf Grund der allmählichen Häufung von Mißbräuchen notwendig erschienen.

Das war beim Computer nicht der Fall – so sehr waren wir von seiner Mystik und Macht bestrickt. Sicherlich haben wir die Leute nicht allgemein mit Computern vertraut gemacht; statt dessen wurden Computer als „Riesengehirne" angepriesen. Die Mystik entwickelte sich zum Autoritätsdenken. Eines der Hauptprobleme der *Autorität* ist, daß man die Schuld auf sie schieben kann. Ich möchte Ihnen ein paar Beispiele zur Kenntnis bringen.

Die Autorität des Computers als Sündenbock und Ausrede

Dies ist wahrscheinlich ein weltweites Phänomen. Man ruft die Firma an, die sich in der Rechnung geirrt hat, die Bank, die widerrufene Schecks nicht retourniert hat, die Vereinigung, die Sie auf die Liste der unverläßlichen Zahler gesetzt hat – und die Antwort: *Tut mir leid, mein Herr, aber wir haben jetzt einen Computer ...*

- Die Allen Piano and Organ Company in Phoenix (USA) brachte in einem Werbespot im Rundfunk, daß ihr Lagerhaltungs-Computer sich bei der Bestellung geirrt hätte; sie hätten nun ein überfülltes Lager und würden deshalb einen Ausverkauf machen. Im Namen der Association for Computing Machinery schrieb ich an die Gesellschaft einen Brief, in welchem ich anbot, den Computer oder das Programm so herzurichten, daß ein solcher Fehler nicht mehr vorkommt. Die Bedingung: *wenn sich herausstellen sollte, daß sich ein Mensch* und nicht der Computer *geirrt hat*, dann müsse die Firma dies in der folgenden Werbesendung zugeben. *Datamation* verfolgte die Angelegenheit, und es stellte sich heraus, daß die Allen Piano and Organ Company *keinen* Computer hatte und auch keine EDV-Einrichtungen benutzte.

  Beachten Sie bitte die feste Überzeugung der Werbeleute, daß ein Computer ihren falschen Angaben Autorität verleihen würde.
- Ich besuchte einmal vier ältere Damen. Als sie erfuhren, daß ich in der Computerbranche arbeite, gab es einen erbosten Chor von Schreckensgeschichten. Dann brachte eine der Damen den Brief ihrer Bank mit einer handschriftlichen Entschuldigung des Kassiers für die Fehler des Computers. Ich saß in der Falle. Um das Gesicht zu wahren, rief ich den Vizepräsidenten der Bank an, um mich zu erkundigen, was getan werden könnte. Sie hatten ebenfalls keinen Computer!

Die Autorität des Computers als Komplize

Der Computer ist ein bequemes Mittel, um Handlungen, die von illegalen bis zu eigennützigen Zwecken reichen, bewußt oder unbewußt, offen oder stillschweigend zu vertuschen.

- Der berüchtigte Equity Funding Skandal [3] wird bestimmt zu einem klassischen Fall, obwohl es einige Zeit dauern wird, bevor herauskommt, wie er begangen wurde. Schon jetzt wissen wir, daß es eine Pyramidenoperation war und daß Computer verwendet wurden, um sowohl Autorität zu verleihen als auch zusätzlichen Schutz vor Entdeckung. Es scheint so, daß etwa 200 Menschen an den Machenschaften beteiligt waren. Dieser Fall könnte als Anstoß für viele regulierende Maßnahmen dienen.
- Die Universität von Michigan besitzt eine Forschungsabteilung, welche die Wirkung verschiedener Entscheidungen und Maßnahmen auf das Bruttosozialprodukt und sein Wachstum untersucht und vorhersagt; dies mit besonderem Bezug auf den Staat Michigan. Die Resultate könnten ohne weiteres in maschinengeschriebenen oder gedruckten Berichten veröffentlicht werden, das ist jedoch nicht der Fall! Ein Computerausdruck begleitet den Bericht, um ihm so *Autorität* zu verleihen. Die Ergebnisse, die ich sah, erschienen mir sowohl falsch als auch irreführend, und sicher hätten das andere auch entdeckt, wenn sie vom Computer ebenso wenig beeindruckt gewesen wären.

Vielleicht kommt einmal der Tag, an dem Amerika seine Umweltschutzbehörde durch eine Menschenschutzbehörde verstärkt. Dann könnte sie nach dem Motto der gegenwärtigen Warnaufdrucke auf Zigarettenpackungen (*„Das Zigarettenrauchen gefährdet Ihre Gesundheit"*) anordnen, daß jeder vom Computer gedruckten Seite folgender Satz vorangestellt werden muß:

„WARNUNG – diese Antworten wurden von einem Computer erzeugt und könnten Ihre Gesundheit gefährden!"

Der Computer als Mistschlucker

Eine bekannte Binsenweisheit der Computer-Verwendung ist „Mist hinein, Mist heraus". Aber was geschieht, wenn wir vollkommen *richtige* Daten eingeben? Können wir sie wieder herausbekommen? Kann das auch jemand anderer tun? Wenn sie herauskommen, sind sie dann auch lesbar und verständlich?

Wir leben noch immer in jenem Zeitalter des Computers, wo 90 oder mehr Prozent der Daten völlig vom dazugehörigen System abhängig sind, das sie in Information umsetzt. Die Datenbeschreibung von COBOL ist ein erster Ansatz, dies zu verbessern, aber warum muß die Beschreibung dem Programm beigefügt sein und nicht den Daten? Braucht man denn ein Programm, wenn man in einer Bibliothek ein Buch lesen will? Die Organisationsmacht des Computers ist völlig von der Lesbarkeit und Austauschbarkeit der Daten abhängig.

Ein klassisches Beispiel ist die Situation, die entstand, als die amerikanische Umweltschutzbehörde durch Konsolidierung mehrerer verschiedener Ämter gebildet wurde, von denen jedes seine eigenen Informationssysteme hatte. Als sie auch die Daten konsolidieren wollten – was sicher einer der Hauptgründe für die Vereinigung dieser Gruppen war – stellte es sich heraus, daß die Daten nicht nur zwischen verschiedenen Komponenten nicht austauschbar waren, sondern nicht einmal zwischen den verschiedenen Computersystemen in den Unterabteilungen der Behörden! Natürlich bewegen sich die Luftmassen über viele Staaten, jeder mit seinen eigenen Computern und Überwachungssystemen, und jeder unfähig, Entscheidungen zu treffen, die für das ganze Land eine Optimierung bewirken würden.

Natürlich könnte dieses Schlucken der Daten weitgehend dadurch gelöst werden, daß Etiketten und Datenbeschreibungen auf Datenträgern verwendet werden, so daß die Daten selbstbeschreibend sind. Der amerikanische Kongreßabgeordnete Brooks hat zu einer *Unabhängigkeitserklärung für Daten* aufgerufen.

Eine andere Möglichkeit, Daten nicht mehr herauszubekommen, ist der Zusammenbruch des Computersystems.

Die Integration von Computersystemen in menschliche Angelegenheiten erfordert extreme Zuverlässigkeit. Das wissen wir alle, und doch gibt es Zeiten, wo wir uns durch die Macht des Computers verleiten lassen, ihm eine Funktion anzuvertrauen, die an bestimmte Stichtage gebunden ist. Selbst ich bin dessen schuldig. Wir verwenden einen Computer für die Textverarbeitung des Honeywell Computer Journals. Das Problem ist, daß wir an einem Computer hängen, der für Software-Experimente und zur Überprüfung von neuen Softwaresystemen oder für Vergleichsläufe in verschiedenen Konfigurationen verwendet wird. Während die Hardware sehr verläßlich sein mag, trifft dies für neuentwickelte Software leider nicht zu – und wir haben unseren gesamten Text den Plattendateien anvertraut. Wenn nun Schwierigkeiten auftreten, gibt es keine manuellen Methoden – wie angestrengt, ja verzweifelt man sich auch bemühen mag –, diese Schwierigkeiten behelfsmäßig zu beheben. Entweder letztlich perfekte Arbeit oder überhaupt nichts – wir sind einem System ausgeliefert, das hundertprozentig funktionieren muß.

Meine Geschichte soll klarmachen, es bedeutet menschliches Versagen, hier optimistisch zu sein und zu glauben, ein Computer könne nicht zusammenbrechen. Dann plant man nämlich weder eine Zweitanlage für den Notfall noch eine Reservedatei auf einem anderen System, noch Stapelverfahren, die auch dann arbeiten, wenn das Timesharingverfahren zusammenbricht. Und dann kann man nicht einmal mehr den Mist verarbeiten.

Einige Maßnahmen zum Schutz des Menschen

Die Gesellschaft hat vor langer Zeit erkannt, daß Tätigkeiten, von denen öffentliche Sicherheit oder Wohlergehen abhängen, gewissen Regulationen und Erziehungs- oder Trainingserfordernissen unterworfen werden müssen. Diese Maßnahmen führten zu Berufsständen mit besonderem Verhaltenskodex und speziellen Arbeitsrichtlinien. Bis jetzt wurden der Computerbranche keine derartigen Maßnahmen auferlegt; man kann nur vermuten, daß die eben erwähnten Berufsstände nicht ebenso plötzlich vor dem gesellschaftlichen Bewußtsein erschienen sind.

Das durch einige Vergehen bekannt gewordene öffentliche Risiko veranlaßte den Staat Kalifornien im Jahre 1971, die Zertifikation von Computerprogrammierern als Berufsgruppe zu erwägen. Die AFIPS gründete im Februar 1972 ein System-Zertifikations-Komitee.

Das Komitee kam sehr bald zu der Erkenntnis, daß es keine Möglichkeit zu geben schien, eine Zertifikation zu erreichen. Ich machte den Vorschlag, die AFIPS sollte eine Reihe von Büchern über Arbeitsrichtlinien und Arbeitspraxis herausgeben. Das erste dieser Bücher über Arbeitspraxis behandelt Vertraulichkeit und Sicherheit, es wird in Kürze in der Praxis erprobt. Es ist weitgehend in Form einer Prüfliste abgefaßt. Das Komitee nennt sich, nebenbei bemerkt, nun *Systemverbesserungs-Komitee* (*„Systems Improvement"*), um zu betonen, daß es nicht der Meinung ist, daß gegenwärtig irgendeine Form von Zertifikation möglich ist.

Zuverlässigkeit für die Integration in menschliche Angelegenheiten

„Reliability for Integration into Human Affairs" war der Titel einer Sitzung der National Computer Conference 1973 in Amerika [5]. Die Sitzung unterschied sich etwas von den anderen. Die anderen Sitzungen sollten *vertikal-* oder *endanwender*-orientiert sein. Als ich das Programm plante, ging ich absichtlich einen Schritt weiter, um zu sehen, welche Aspekte des Computersystem-Entwurfs vielen Endanwendungen gemeinsam sind, aus dem spezifischen Grund, daß sie direkt in menschliche Angelegenheiten integriert waren.

An dieser Sitzung nahmen Vertreter von Flugüberwachungssystemen, Patientenüberwachung durch den Computer, Kraftwerkssteuerung durch den Computer, Kreditsystemen, Transportunternehmen und des Handels teil. Viele dieser Anwendungen gehören zu Kategorie 3; die Verbundnetz-Steuerung erfordert zum Beispiel gegen einen Netzzusammenbruch eine Reaktion, die viel schneller erfolgen muß, als ein Mensch dies könnte. Flugüberwachung ist ein weiteres Beispiel; man erwartet, daß in den achtziger Jahren ständig 5000 Menschen über Los Angeles in der Luft sein werden, in 700 Flugzeugen! Der Sprecher legte zwei wichtige Forderungen vor:

- Die voraussagbare Zuverlässigkeit sollte astronomisch sein.
- Es sollte eine Sicherheitsvorrichtung vorhanden sein für den Fall, daß das System ohne Vorwarnung zusammenbricht.

Auf der Suche nach anderen Elementen für gutes Verhalten stellte sich heraus, daß keinem der Diskussionsteilnehmer oder ihren Entwurfsteams Quellen oder Nachschlagewerke bekannt waren, die sie für die Zuverlässigkeitsaspekte der Computeranwendung verwenden konnten, obwohl ihre Anwendungen viel Gemeinsames aufwiesen. *Bob Patrick* zitiert einige Beispiele für schlechten Entwurf:

- Ein Rechenzentrum hielt die Zweitbänder in einem feuersicheren Tresor und die „Großvater-Bänder" im Inneren eines Berges. Es gab jedoch nur eine Kopie des „Log-Buchs" (Operating-Protokolls), das dem Operator sagt, wie die Bänder zu lesen sind, und dieses befand sich im Maschinenraum, wo es bei einem Brand zerstört würde.
- Eine militärische Einrichtung hatte einen hohen Sicherheitsfaktor und schützte ihre Daten sehr gut. Bei der Neuaufnahme von Daten beschnitt man periodisch die ersten 20 Meter der Bänder. Das Problem war, daß diese Bandstreifen in den Müll wanderten. Die Bänder waren nicht gelöscht!

Donn Parker, der Vorsitzende der erwähnten Sitzung, ist ein anerkannter Fachmann für Computer-Verbrechen. Er schätzt, daß der Schaden gegenwärtig jährlich 300 Millionen Dollar beträgt und bis 1980 bereits 2 Milliarden Dollar erreichen wird! *Dick Mills* von der First National City Bank sagt, daß durch die „digitalen Kanäle" der Bank täglich 8 Milliarden Dollar laufen, so daß auch durch eine kleine undichte Stelle eine Menge Geld verlorengeht. Ich glaube nicht, daß wir übervorsichtig sind, wenn wir auf extremer Zuverlässigkeit bei derartigen „menschen-sensitiven" Anwendungen bestehen.

Schutzmaßnahmen gegen Mißbrauch

Gesetzliche Maßnahmen

Es gibt viele Beispiele von Gesetzen zum Schutz der Person. Bauarbeiter müssen Helme, Motorradfahrer müssen Lederkleidung und Helme tragen. Das sind Schutzmaßnahmen im Beruf, die dem einzelnen aufgezwungen werden, vermutlich, weil er eine von der Gesellschaft geleistete Investition darstellt.

Die amerikanische Regierung hat für die Herstellung von Autos bestimmte Sicherheitsbedingungen erlassen. Es scheint sicher, daß der Computer nicht nur direkte Auswirkungen auf die Sicherheit unserer Bürger hat, sondern auch auf andere Rechte. Es muß daher vernünftig sein zu verlangen, daß Software und Hardware ebenfalls nach bestimmten Normen gebaut werden, um diese Rechte zu schützen. Dieses Gebiet wird umfassend in *Juristische Aspekte von elektronischen Informationssystemen* [6] behandelt.

Freiwillige Maßnahmen

Dr. *Harold Sackman*, Vorsitzender des AFIPS Komitees über die Auswirkungen des Computers auf die Gesellschaft, verlangte vor kurzem eine *Computeranwender-Gesellschaft für Amerika*. Das sollte eine Gruppe von Computeranwendern sein, die sich aktiv mit sozialer Zuverlässigkeit befaßt, und zwar deshalb, weil die EDV-Gemeinschaft die Probleme wirklich als erste zu sehen bekommt und dafür verantwortlich ist, die Probleme jenen aufzuzeigen, die sie behandeln können. Die ACM verdankt der skandinavischen Erfindung des Ombudsmannes viel; ihr Ombudsmann-Programm hat viele Probleme schlechter Computeranwendung gelöst.

Will man die elektronische Datenverarbeitung zu einem echten Beruf machen, so sind Handbücher für Entwurf und Praxis erforderlich. Viele Computergesellschaften verwenden mehr oder weniger entwickelte Praxiscodes und Zertifikationscodes für EDV-Leute. Man hofft, daß sie nicht vor

einer allgemeinen Zertifikation zurückschrecken, sondern auch eine anwenderorientierte Zertifikation in ihr Programm aufnehmen, und zwar in Zusammenarbeit mit den betreffenden Berufsgruppen. Wir werden unsere Systeme mit Leistungsmessungs- und Auswertungseinrichtungen versehen müssen.

Schlußfolgerung

Als Hüter der Macht tragen wir viel Verantwortung. Als ich die ACM-Tagung 1970 plante, geschah dies als Modell für ein Nationales Computerjahr, welchem möglicherweise ein Internationales Computerjahr folgen könnte [7]. Die Liste der Ziele für ein solches Jahr könnte so aussehen:

- Eine bewußte Indienststellung von Computern für internationale Ziele, eine Stärkung des öffentlichen Verständnisses für die Rolle und die Möglichkeiten der Computeranwendung sowie eine Betonung der Rolle des Computers als Diener durch stärkere Humanisierung von Anwendungen und Verwendung.
- Die Entwicklung von Zukunftsstrategien für die beste Verwendung von Computersystemen (technologisch, gesellschaftlich, erzieherisch, politisch und legislativ).
- Konservierung und Maximierung der existierenden und zukünftigen intellektuellen Ressourcen – bekannt als Daten und Programme –, indem man herausfindet, wie sie auf multiplen Geräten und in multiplen Anwendungen verwendet werden.
- Hilfestellung für die Regierung, Industrie und die private Entscheidungsfindung durch Erschließung neuer und vollständigerer Daten für diese Entscheidungen. Erleichterung dieser Entscheidungsfindung durch Verminderung des erforderlichen Informationsvolumens (im Gegensatz zu Datenvolumen).
- Planung eines geschlossenen Zyklus zur Neuaufteilung von Arbeitsprozessen zwischen Menschen und Computern, zur Neueinschulung vor der Änderung des Arbeitsgebietes, so daß die Menschen ihre Aufgaben optimal erfüllen können.
- Garantie einer entsprechenden Rücksichtnahme auf öffentliche Sicherheit und Gemeinwohl, dann wenn Computer direkt in den menschlichen Bereich integriert werden.
- Erstellung neuer und breiter interdisziplinärer Kanäle für den Informationsaustausch zwischen bisher getrennten Organisationen; Förderung ihrer Beteiligung auf internationaler Ebene.
- Planung der wirtschaftlichsten und effektivsten Wechselwirkung zwischen Computersystemen und anderen Systemen, z. B. Kommunikationssystemen.

Es ist nicht zu früh für eine umfassende Untersuchung der Wechselwirkung zwischen Computern und unserer Gesellschaft. Zwei Beiträge zur National Computer Conference 1973 [8] erhärten diese Aussage – *Die sozialen Auswirkungen der Computeranwendung über nationale Grenzen* und *Ein neuer Vorstoß der NSF – Einfluß des Computers auf die Gesellschaft* (NSF = National Science Foundation of America).

Ich habe das Gefühl, daß es für Computer und Gesellschaft nicht so schwierig sein wird, sich einander anzupassen, wenn wir uns wirklich bemühen. 1970 sagte der stellvertretende Postminister der USA, daß ein Drittel aller Poststücke maschinell adressiert werden, jedoch nur 6% kommen am Postamt mit der richtigen Postleitzahl an. Er fragte, warum die elektronischen Adreßdateien nicht ebenso nach Postleitzahlen geordnet werden können wie nach anderen Gesichtspunkten? Ich stellte dann diese Frage vielen Datenverarbeitungsabteilungen. Die Antwort war, daß sie daran nicht gedacht hätten und daß sie genausogut dieses Ordnungskriterium verwenden könnten.

So einfach ist das.

Literatur

[1] *Yoder, R. D.:* Management of Computer Failures in Clinical Care. – Datamation 18 (1972), No. 10, pp. 78–82.

[2] *Bruck, H.:* To Redress the Balance. – In: R. W. Bemer (Ed.) Computers and Crisis. Proceedings of the 1970 ACM Annual Conference, ACM, NY 1971.

[3] *McLaughlin, R. A.:* Equity Funding: Everyone is pointing at the computer. Datamation 19 (1973), No. 6, 88–91.
N.N.: This Swindle Could Happen Here. The Economist 247:10 (1973) April 14.

[4] *Richardson, J.,* in *R. W. Bemer* (Ed.): Computers and Crisis. Proceedings of the 1970 ACM Annual Conference, ACM, NY 1971.

[5] *Parker, D.:* Reliability for Integration into Human Affairs, Session at the AFIPS National Computer Conference 1973.

[6] Legal Aspects of Computerized Information Systems, US Government Report. Reprinted in: Honeywell Computer Journal 7 (1973), No. 1.
Ebenfalls erhältlich als „Cosati 73–01" vom National Technical Information Service (NTIS), Springfield VI 22151.

[7] *Bemer, R. W.* (Ed.): Computers and Crisis. Proceedings of the 1970 ACM Annual Conference, ACM, NY 1971.

[8] *Nanus, B., Wooton, M., and Borko, H.:* The Social Implications of the Use of Computers Across National Boundaries. Proceedings of the AFIPS National Computer Conference 1973, AFIPS Press, NJ 1973, pp. 735–745.
Lykos, P.: A New NSF Thrust – Computer Impact on Society. Proceedings of the AFIPS National Computer Conference 1973, AFIPS Press, NJ 1973, pp. 747–749.

METRIC MEETING

THE U.S. METRIC ASSOCIATION, SOUTH BAY AREA CHAPTER PRESENTS A JOINT MEETING WITH LOCAL CHAPTERS OF

ACM, ASQC, CSPE/NSPE, NCSL & PMA

DATE: THURSDAY, JUNE 16, 1977

LOCATION: PROUD BIRD RESTAURANT "ESCADRILLE" CONFERNCE ROOM
11022 AVIATION BLVD, LOS ANGELES, CA. (213) 670-3093

KEYNOTE SPEAKER: DR. GEORGE FISCHBECK, KABC-TV CHANNEL 7

FEATURED SPEAKERS

HARRY KINNEY
Mayor of Albuquerque, N.M.
Presidential Appointee to the
U.S. (Federal) Metric Board

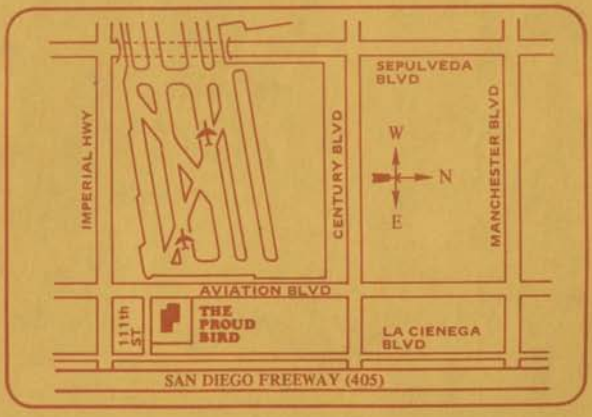
ROBERT BEMER
Consulting Engineer
Honeywell Information Center
Phoenix, Arizona

MODERATOR: EARL G. MILLS, CHAIRMAN, METRIC BOARD, HUGHES AIRCRAFT CO.

AGENDA:
5:30 to 6:30 – Social Hour
6:30 to 7:30 – Dinner, Keynote Speaker Dr. George Fischbeck
7:30 to 8:15 – Harry Kinney
8:15 to 9:00 – Robert Bemer

DINNER: Reserved $7.00,
Without Reservations $8.00

RESERVATION BEFORE JUNE 8 BY CHECK TO: DR. ALBERT P. MATTAY, 261 S. CANNON DRIVE, BEVERLY HILLS, CA. 90212. MAKE CHECKS PAYABLE TO: SOUTH BAY AREA CHAPTER OF USMA – PHONE: (213) 535-1933 TRW DSSG



DINNER OPTIONAL • DOOR PRIZES • EVERYBODY WELCOME

acm
Association for Computing Machinery
Los Angeles Chapter

DATA-LINK

VOLUME XXI, NO. 10 JUNE 1977

JOINT L.A. CHAPTER/L.A. SIGSPACE DINNER MEETING

Wednesday, June 1, 1978

BUILDING A SELF-FIXING SPACE-BORNE COMPUTER

PROGRAM

The U.S. Air Force is developing fault tolerant computer architecture and software to assure system reliability to support mission operations of five or more years. A system brassboard has been developed and funding has been substantially increased to provide for the requirements of our future space-craft needs. The objectives of the fault tolerant computers is to have a five year operational probability of 95%. Captain Kern will discuss the background of this project and the progress which has been made to date.

SPEAKER

LARRY KERN

Captain Larry Kern, under the Director of Technology SAMSO Lt. Col. Frank Dyke, has been responsible for the fault tolerant computer program. He is the principal architect of what has become a top priority technology project in the U.S. Air Force. For his work in this area, Captain Kern has received the Obenland Engineering Award. Captain Kern received his BSEE degree from San Jose State University.

PLACE

Cockatoo Restaurant
4334 West Imperial Highway
Hawthorne, California

TIME

Executive Council	5:00 p.m.
Cocktails	6:00 p.m.
Dinner	7:00 p.m.
Program	8:00 p.m.

MENU

Fruit Cup
Eye of Prime Rib, Sauce Diablo
Anna Potatoes
Broiled Tomato
Apple Pie

PRICE

Dinner:
Tax and tip included
WITH RESERVATIONS $6.50
Program only Free

RESERVATIONS

Robert Lewis (213) 481-3368
before 10:00 a.m., Tuesday, May 31, 1977
Please note that this is a new phone number

IN THIS ISSUE . . .

Report of the nominating committee 2
L.A. SIGBDP plans two summer tours 2
New L.A. ACM members 2
ULTC celebrates nineth year 2
L.A. SIGGRAPH plans summer meetings 2
Going metric - joint meeting announcement 2
The economics of technology 3
Thanks to our DATA-LINK sponsors 3
San Fernando Valley Chapter announcement 3
SIG notices 4
Calendar of events 5
Managing data processing in the distributed environment 6
Combating the blame-the-computer syndrome 6
Information flow and the micro image 7
The privacy report is still available 7

REPORT OF THE NOMINATING COMMITTEE

The following persons have been elected to the L.A. ACM offices indeicated for the term commencing July 1, 1977 and ending June 30, 1978.

For Chairman	Mike Ikezawa
For Vice Chairman	Gloria Silvern
For Secretary	Michael Walsh
For Treasurer	Joan King
For Records Controller	Mary Brennan
For Members-at-Large (2)	Marsha Hopwood Frank Ingrassia

L.A. SIGBDP PLANS TWO SUMMER TOURS

As in the past, L.A. SIGBDP is planning on holding tours of interesting computer locations during the summer. No specific plans have been announced, but all L.A. ACM members will receive information before each event.

NEW L.A. ACM MEMBERS

Carl Abe	Los Angeles Police Department
Prof. David Farber	UC Irvine
Frank Frazier Jr.	TRW
John Grayson	GTE Data Services
Russell Harrell	T&AM Agency
Mark Helfen	JPL
Jose Icaza	Wordplex Corporation
Miles Lawrence	Ralph M. Parsons Company
Guy Lohman	JPL
Diane Millard	TRW
Bob Patrick	self-employed
Mike Rogson	RAND Corporation

ULTC CELEBRATES NINETH YEAR

The Los Angeles Urban League Training Center, in operation since September 1968, was organized jointly by the Bank of America Foundation, IBM Corporation and the Greater Los Angeles Urban League. Its purpose is to teach qualified men and women, particularly those who are unemployed or under-employed, marketable data processing skills, and to place them in the computer field.

In its nine years of operation, the center has placed more than 1,000 graduates in jobs, mostly in Southern California. They have a high, 71 per cent job retention rate.

In addition to classes taught at the center this year, the computer is supporting a programming class at United Community Efforts, a vocational center in East Los Angeles. Several terminals at UCE, connected to the ULTC computer, are used to help teach COBOL programming. Next month UCE will expand its offerings, starting a new data entry class incorporating key-to-disk, keypunch and CRT terminal skills.

ULTC graduates represent COBOL programming, computer operations, System 3/ programming and operations, as well as clerical skills. If your company is interested in learning more about ULTC graduates, please contact Sandra Carter at 753-4244.

L.A. SIGGRAPH PLANS SUMMER MEETINGS

This summer, L.A. SIGGRAPH plans to hold at least one meeting. The meeting currently planned is for Tuesday, July 12 at 8:00 p.m. The topic to be discussed is "A review of the current state of the art of military command and control display systems." No speaker, or meeting place has been firmed up yet. For more details, contact Jerry Johnson at 825-6948.

THE U.S. METRIC ASSOCIATION, SOUTH BAY AREA CHAPTER, PRESENTS A JOINT MEETING WITH LOCAL CHAPTERS OF

ACM, SIGMETRICS, ASQC, CSPE/NSPE, NCSL & PMA

GOING METRIC

KEYNOTE SPEAKER: George Fischbeck, KABC-TV Channel 7

FEATURED SPEAKERS: Harry Kinney, Mayor of Albuquerque, NM
Presidential Appointee to the U.S. (Federal) Metric Board

Robert Bemer, Consulting Engineer
Honeywell Information Center
Pheonix, AZ

ABSTRACT: Mr. Kinney will speak about the various aspects of metrication on city, state, and government levels. Mr. Bemer will speak about the metric changeover and how it is and will affect the computer business and software industry.

The after-dinner program is free, and there will be door prize tickets on sale. We heartily encourage all L.A. ACM and SIGMETRICS members to attend this special meeting.

DATE: Thursday, June 16, 1977

TIME: 5:30 p.m., Social hour
6:30 p.m., Dinner and keynote speaker
7:30 p.m., Featured speakers

PLACE: Proud Bird Restaurant, Escadrille Conference Room
11022 Aviation Blvd. (between Imperial Hwy and Century Blvd., at the end of LAX runways)

PRICE: $7.00 for dinner with mailed reservation
$8.00 for dinner without reservation

RESERVATIONS: Mail reservation and payment by June 8 to:
Dr. Albert P. Mattay
261 S. Cannon Dr.
Beverly Hills, CA 90212
(make check payable to South Bay USMA)

Computer History Museum

SEPARATION SHEET
Computer History Museum

Category of item(s):

- ☐ Audiocassette
- ☐ Electronic records
- ☐ Oversized materials
- ☒ Photographs/negatives
- ☐ Printed materials
- ☐ Three-dimensional objects
- ☐ Videotape
- ☐ Other: _____

Number of items: 79 35mm slides (4 slide sheets)

Provide information about item(s):

Presentation slides about converting to the metric system. Found in the 1977 part of the folder, so likely from 1977.

Originally filed in:

Catalog #: 102785430 Lot #: X3054.2005

Collection Name: Robert (Bob) Bemer papers

Series: Papers Subseries: _____

Folder title: Speeches and papers, 1955-1982
Folder 4 of 5

Relocated to:

Catalog #: 102785615 Lot #: X3054.2005

Collection Name: Robert (Bob) Bemer papers

Series: Papers Subseries: _____

Folder title: Presentation slides, 1970-1978

Purpose of removal: To separate the slides from the text.

Date: 2/28/19

Name: Sydney Olson

RWB COPY

Yes, It's Inevitable -- and You WILL Enjoy It!

R. W. Bemer, Honeywell Information Systems
Phoenix, AZ 85005, US

We've been forced (economically, by the Common market) to convert to the metric system. SI, that is. With the change come trauma and expectation of heavy expenses. Relax. It's going to be fun when you see how the metric system will make many things so easy, including thinking. And there will be big opportunities for actually saving more than conversion costs.

Now an avid supporter of the SI (Système International), I've changed measurement systems twice, because the SI is not the CGS metric system that I learned first. But I understand people's reluctance to learn a new system when they have the old one so well in hand, and use it so effectively.

For example, the matter of expressing volume. The SI has only the litre and cubic metre, whereas our old system has a profusion. In fact, the historical reason for having several units may be the capability to express volume in units of one. Instead of saying so many litres, we can just say a firkin, a hogshead, a barrel, or one barrel. And everyone knows that a barrel contains 31.5 gallons.

SLIDE (3) - 31.5

Except if it contains petroleum products, in which case it's 42 gallons.

SLIDE (4) - 42

And except for malt beverages, in which case it's 31 gallons.

SLIDE (5) - 31

Of course these barrels are for liquids, but the barrel is also defined for dry content, measured in quarts rather than gallons. Everyone knows that a barrel for dry content contains 105 quarts.

SLIDE (6) - 105

Except if it contains cranberries, in which case it's 87 quarts.

SLIDE (7) - 87

It can be interesting to see the correspondence, so we convert according to official definition:

SLIDE (8) - conversion

A dry quart has the same volume as 1.1636 liquid quarts. Calculating 105 times 1.1636,

SLIDE (9)

divided by 4 quarts to the gallon, we apparently discover a new barrel, of 30.54 gallons. Let's call it the "barrel for dry liquids"!

SLIDE (10) - "dry liquids"

You may think that this is a variation in one uncommon unit, and that such inconsistencies do not exist in, say, linear measure. Let's look for the lowest common denominator in linear measure. In doing so, we observe that the old system is again rich in terms, whereas the metric system user has only the metre.

SLIDE (11) - Lowest Common Denominator

1 statute mile	=	4	quarters
	=	8	furlongs
	=	80	chains (survey)
	=	176	chains (football)
	=	320	rods
	=	880	fathoms
	=	1 760	yards
	=	5 280	feet
	=	7 040	spans
	=	8 000	links
	=	15 840	hands
	=	21 120	palms
	=	28 160	nails or ells
	=	63 360	inches
	=	70 400	lines (buttons)
	=	84 480	digits
	=	190 080	barleycorns
	=	760 320	lines
	=	3 041 280	irons
	=	4 055 040	ounces
	=	4 561 920	points
	=	63 360 000	mils

You think I'm mistaken in showing the number of ounces per mile? Not so. The ounce is equivalent to 1/64 of an inch, and is used in measuring leather thickness!

When one thinks he has found the lowest common denominator for this set, there will always be another to confound it. The "engineer's chain", for example, is either 50 feet or 100 feet. And I have'nt even introduced all units of the textile industry -- spindle, hank, heer, cut, skein, lea, bolt, roll, and thread. 220 spindles are 1890 miles long!

But the system is obviously rich in relationships. For example:

SLIDE (12)

If the average horse is 15.84 hands tall
(and we know that 1 mile = 15 840 hands),
does 1 mile = 1000 horses?

And that's for just the statute mile. How about sea horses and the nautical mile?

Even the foot is subject to variance. Perhaps the Philadelphia foot is the strangest of all.

SLIDE (13) - PHL

It's 1.002 statute feet in the East-West direction, and 1.0045 statute feet North-to-South! This is nontrivial, for it pertains to the city maps that tell you, for example, where the sewer pipes and the telephone cables are buried!

That's enough about the old system; you can find many more oddities yourself. But imagine the time wasted in learning these relationships, and especially the computational difficulties that may arise!

The SI Aids Learning

The metric system should help everyone in learning mathematics, engineering, and other quantitative disciplines. The concepts of both number base and exponentiation are built into the system, in the mechanism of prefixes.

SLIDE (14) - Metric Prefixes

10 to +i	i	10 to -i
exa (E)	18	atto (a)
peta (P)	15	femto (f)
tera (T)	12	pico (p)
giga (G)	9	nano (n)
mega (M)	6	micro (μ)
kilo (k)	3	milli (m)
hecto (h)	2	centi (c)
deka (da)	1	deci (d)

"Exa" and "peta" may be new to some. They were adopted quite recently, to fill out the range. And what a range it is! Sufficient for all galactic measurement, and all the way down to subatomic measurement.

Observe "nano", which is only 10 to the -9th. Do you know how far light and electricity travel in a nanosecond? Let's calculate it. The speed of light is very close to

SLIDE (15) - 300 Mm/s (megametres/second)

We see "nano-" defined as 10 to the -9, and "mega-" as 10 to the 6.

SLIDE (16) - (prefixes again)

Because 6 - 9 = -3, which we see from the table is "milli-", the speed of light equals

SLIDE (15) - (back up to speed of light)

300 mm/ns (millimetres/nanosecond)

Simple, isn't it? Light travels 300 millimetres in a nanosecond -- just 5 mm short of the old foot.

If you wonder why this example was chosen, remember that computers now operate at nanosecond speeds and less, so it's a tough problem to get the results anywhere by wires that must be under 300 mm in length. Despite microcomputers, we still have some computers and peripherals that fill a good-sized room.

Of course, as many people have hand calculators today, we could have done it in the old English system by multiplying 186,283 by 5280, watching the decimal point, then shifting it left 9 places. But it's not as easy as the metric way.

SLIDE (16) - (forward to prefixes again)

Note that above 3 there are no powers except multiples of 3. This practice breeds better comprehension, like marking off three's in writing numbers of many digits. Note also, as a memory convenience, that all symbols are upper case for powers greater than +3.

SLIDE (17) - (distraction)

The SI Abhors Fractions

How about those fractions? The SI doesn't have fractions as such. Do many people have hand calculators that work in fractions? So they can do problems like:

> If Sam can dig 32 ditches in 9 hours, and Joe can dig 24 ditches in 7 hours, how long will it take them to dig one ditch - together? If they can keep out of each other's way?

I have often wondered whether such problems have much real applicability -- if perhaps they just were dreamed up to teach the difficult art of fractions. Fractions may not be difficult to you, but they are to the French, who cannot understand where that 0.375 inch dimension came from, because 3/8ths means nothing at all to them. My company had to import teaching materials from Italy to show the French why such numbers existed, because it is illegal to sell nonmetric materials in France.

The only fraction the French use is 1/2. They will order a demilitre of wine, but a smaller amount is not a quarter litre, which would be 2.5 decilitres, but rather an even 2 decilitres. So you see that our fractional system is the awkward octal or hexadecimal one of computers. Theirs goes 2, 1, 0.5, 0.2, 0.1, 0.05, etc., to form a sequence that repeats in 10's.

And fractions are difficult to my wife, because I have to scale our recipes up and down, and I think difficult to most people. Ask yourself what half of 5 5/8 is. Got it? It's actually a very useful figure.

SLIDE (18) - 2x6

You would use it to saw a 2x6 in half, because the rough size of 6 inches becomes 5 5/8 inches when finished. How much easier to ask "What's half of 140?"

The Why of Converting to the SI

We'll come back again to the value of the metric system to the learning and calculating processes, after a digression to question why the United States is converting at just this time. After all, the metric system has been our only legal system for over 100 years, and all of our customary units are defined by their metric values to avoid chaos.

SLIDE (19) - Cost too much?

Legality is apparently not sufficient to avoid a few minor problems. For example, I have my mass as 83 kilos and my height as 185 cm on my driver's license. The last time I was checked by a highway patrolman, he looked at the license and asked how tall I was. I said 185, and he said 185 what? I said centimetres. He said "Come on, now, what's your real height?

So could it be that we have finally realized that the SI, the metric system, is really so much better and easier that we should convert now despite the possibly high cost of that change?

Not in this democracy! The real reason is:

SLIDE (20) -

1978 January 01

The date when non-metric products
may not be sold in Europe!

The US Government may be getting around to it only now, but General Motors, IBM, Ford, Honeywell, Rockwell, and hundreds of other companies have been converting for a long time. The reason is plain. Trade, and the balance of payments factor, which translates directly into jobs and well-being for the American people.

Kenyon Taylor of the U.S. Metric Conversion Board, appearing on the TODAY show in New York, said that the loss of export if we didn't convert to metric design would amount to about 5 million jobs!

General Motors is very serious about export.

SLIDE (21) - GM plastic a'card

You might as well get used to the unit "kilopascal", because that's how we'll measure the pressure in your tires. Don't worry about the gas pump reading litres -- the dollar amount is the thing that counts.

Conversion and Dollars

General Motors has one answer to the question "Doesn't it cost too much to change?"

SLIDE (22) - Fanbelts

They now have over 950 different fanbelt sizes that they (and service stations) must stock. When fully converted to metric manufacture, there will be fewer than 50 different sizes! Here's a clear case where a profit is turned on conversion, not a loss. Your garage will be more likely to have your size in stock if the required inventory is smaller. Don't forget, over 40% of the cars on U.S. roads today are metric.

15 MIN

In many other cases such as this, forced redesign has yielded economic benefits that were lurking there all the time, if we hadn't been so blind. They may often exceed conversion costs. In manufacturing, where most conversion costs were thought to lie, ways have been found to minimize them, and spread them over a period of years.

Conversion and People

The impact upon people is a larger problem. There are many diehards who will fight the SI. The best way to rout them is to avoid conversion entirely (because it's a crutch), and particularly to avoid exact conversion. Most exact conversion is not justified because of inherent tolerances.

SLIDE (23) - 2" (maybe) nail

Those nails are just as likely to be chopped off at exactly 50 millimetres as they are at exactly 2 inches.

The diehards often use exact conversion to deride the metric system. Somehow football seems to be a favorite target:

SLIDE (24) -

"Bones made the 9.144
metres for first down"

But they don't mind singing

SLIDE (25) -

"I love you --

2219.91 cubic inches and
554.84 cubic inches"

Which translates to:

SLIDE (26) -

"I love you --

a bushel and a peck"

Whomever they love that much, that 2774.2 cubic inches of "you" (assuming they just barely float in water) comes to just 100 pounds -- which is OK if their love is 5' 3".

The diehards don't have to worry about football. That is one sport that will never be converted, even though the Canadian football field is 100 metres long. All U.S. stadiums have the stands crowding the 100 yards and goal areas as closely as possible, and it would be prohibitively expensive to change. Your children, and their children, will still play football, and they will just think of the "yard" as an archaic term peculiar to football, just as you and I think of the "cubit" as an archaic term peculiar to ark-building.

During the course of legislation, conversion was raised as a bugaboo by various union groups. The carpenters, for example, wanted the US to give each of them a huge allowance to buy new tools. This prompted the NBS to make up some displays like this:

SLIDE 27 - Metric and Oldfashioned Hammers

SLIDE 28 - Metric Plane

Of course, a journeyman carpenter starts out with tools worth much less than the allowance they requested.

The electricians couldn't be too recalcitrant about converting to the SI. They've been using it right along!

SLIDE 29 - Metric light bulb

There will be those that feel there are too many new things to learn in the metric system -- particularly the prefixes. Here are some words starting with "P", as taken from the Style Manual of the US Government Printing Office, which is still hopelessly out of date with respect to the SI.

SLIDE 30 -

Which of these English words
is a unit of length/distance?

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometer	prisometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Can't you just hear the man on the street saying

SLIDE 31 - NO TALK

"Is a parameter more or less than a meter?"
"How many perimeters in a parameter?"

And according to the formation rules of the SI he would be pronouncing it: "Is a para meter more or less than a meter?" "How many peri meters in a para meter?" After all, the radio announcer doesn't say "ki-LAW-hertz", and the scientists don't say "ki-LAW-grams". So "ki-LAW-meter" is wrong, too.

Wouldn't it be easier this way?

SLIDE (34) -

Is it any easier now?

20 MIN

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometre	prisometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Now do you spot the "picometre", which is 1/100th of an angstrom?

SLIDE (33) - micrometer

That's pronounced "A MI-cro-ME-tre is so small that it takes a mi-CROM-e-ter to measure it". This difference in pronunciation carries over into hyphenation, so you can tell which is which when reading. Moreover, the international spelling M.E.T.R.E distinguishes between instruments and such, and the unit of length or distance. Handy, isn't it?

The major argument for the "ER" was phonetics. But English is not a phonetic language, and the English spelling is "RE". German is a phonetic language, and it is true that they spell the unit of length "METER". But they don't have our problem! For the instrument, they use the word "Zahler", meaning "counter". In fact, Norway is now converting to the "RE" spelling, even though they have used the metric system for many years.

Can you imagine the assembly line worker at American Motors, putting the 4.0 litre hardware on Gremlins? How will he know to put on the "4.0 litter" for the US, and the "4.0 litre" for export?

TEXTBOOKS

The SI is Coherent

The chief pedagogical value of the metric system is its coherency.

SLIDE (34) - Australian Chart (just for directed graph coherency)

SLIDE (35) - Cube for length-volume-mass

A cube one Benson & Hedges cigarette (100 mm) on a side has a capacity of one litre; the water required to fill it has a mass of one kilogram.

While living in Paris, I found coherency in my hat size. In the U.S. it's 7 5/8, a fractional value that I cannot relate to anything at all. I wrapped a tape measure around my head, read 61 centimetres, ordered a size 61 beret at Galeries Lafayette, and -- voila! -- it fit!

The coherency will also be found in the kitchen, where the confusing pounds, quarts, cups, tablespoons, and teaspoons disappear, to be replaced by millilitres. You won't notice the difference between the 4.9 millilitre teaspoon and the 5 millilitre measure that will replace it. I had to memorize that there were 3 teaspoons per tablespoon, and I still get unsure at times. No memorization is needed when the units are 5 ml and 15 ml. In Europe, it is common to combine ingredients with a mass scale. Here is a metric recipe for your inspection, if not cooking:

SLIDE (36) - Metric cupcakes GRAMMES

Let's look at the metre stick, which is usually marked off in centimetres. Imagine it's a dollar in length.

SLIDE (37) - Metre stick

Halfway for a dollar is 50 cents. Halfway for a metre is 50 centimetres. I've noticed that people learn particularly well when you relate things to money. And in this coherent system the same scale can be used to represent temperature (in degrees Celsius).

SLIDE (38) -

Think of it as water freezing at stony broke, boiling at a dollar. Our body temperature is 37 cents, and the thermostat should now be set to 18 cents or less. Again, the value ratio of 18 cents to the dollar is well understood, and so one gets the "feel" of Celsius temperature. Never convert!

Note the little arrow, sitting uncaptioned at 46 degrees. That's the hottest it gets in my home town (Phoenix) in the summertime. On a trip to Stockholm, I took Icelandic Airlines and made the overnight stop there. On the way to Keflavik airport in the bus the next morning, I sat next to a young man, whom I queried as to the temperature. He replied "5 degrees". I remarked that it was 46 when I left home yesterday. ,space SLIDE 39 - Iceland (N pole)

When I assured him that it was indeed Celsius, his eyes became very large. I suppose such a temperature to an Icelander would remind him more of their volcanoes.

The same scale also works for understanding "mass".

SLIDE (40) -

SLIDE (41) - Keyboard (starting computer part)

Computers and the SI

We should see some extensive modification of computer usage in the transition to the SI. The effect upon hardware will be slight, mainly in the output display area. Standard ISO 2955, "Information Processing -- Representations of SI and other units for use in systems with limited character sets", gives replacement symbols, mostly to get around the upper-case-only problem. But most new devices are dual-case now, and at this moment there is a second-draft proposed 8-bit ASCII standard in circulation (X3L2/77/08); in particular, it has the "mu" for micro-, "omega" for ohm, and the degree symbol.

We expect little trouble with Numerical Control. The APT language is already provided with the necessary metric specification capability. The three areas of software impact are conversion and learning, databases, and program rewrite to take advantage of SI coherency to reduce running costs.

The NBS provides a computer program to convert from customary to SI units, but I expect that it will be superseded by programs of the type I shall describe now -- suitable for both conversion and training.

Here are some possibilities in CAL (Computer-Assisted Learning) programs. Imagine dialing your friendly computing service from your terminal.

SLIDES 42 - 48

```
Term is?  x/y
Retry.

Term is?  a/bq
 - a quantity of no meaning in the SI
```

```
Term is?  m/s
output =    m/s
 - a quantity of velocity

Term is?  M/MS
output =   um/s
 - a quantity of velocity
```

```
Term is?  meter/sec
A "meter" is an instrument.
The unit of length is the "metre".
output =    m/s
 - a quantity of velocity
```

```
Term is?  mm per s.
output =   mm/s
 - a quantity of velocity

Term is?  cm/nanosec.
output =  10 Mm/s
 - a quantity of velocity
```

```
Term is?  kg.m per s**2
output =    N
 - a quantity of force, the "newton"

Term is?  MN.m/km**3
output =   kJ/m**3
 - a quantity of energy density
```

```
Term is?  n.m/s
output =    W
 - a quantity of power, the "watt"

Term is?  newton-metres per seconds
output =    W
 - a quantity of power, the "watt"
```

```
Term is?  n.m/k
output =    J/K
 - a quantity of entropy

Term is?  N.M/K
output =    J/K
 - a quantity of entropy
```

Although only the output is shown here, a simple option can be called to also display input mistakes, if any.

SLIDES 49 - 65

```
Term is?  W/M.k
output =     W/(m.K)
 - a quantity of thermal conductivity

Term is?  n.m/KG.k
output =     J/(kg.K)
 - a quantity of specific entropy

Term is?  a.S per v
output =     F
 - a quantity of electrical capacitance, the "farad"

Term is?  v/a
output =     ohm
 - a quantity of electric resistance, the "ohm"

Term is?  vs/amperes
Retry.

Term is?  volts/amps
output =     ohm
 - a quantity of electric resistance, the "ohm"

Term is?  w/a
output =     V
 - a quantity of electromotive force, the "volt"

Term is?  v.s/a
output =     H
 - a quantity of electric inductance, the "henry"

Term is?  V/m
output =     V/m
 - a quantity of electric field strength

Term is?  W/SR
output =     W/sr
 - a quantity of radiant intensity

Term is?  lumens per square metre
output =     lx
 - a quantity of illuminance, the "lux"
```

```
Term is?  cd.sr/square metres
output =     lx
 - a quantity of illuminance, the "lux"

Term is?  candelas/square m
output =     cd/m**2
 - a quantity of luminance

Term is?  candlepower per sq m
output =     cd/m**2
 - a quantity of luminance

Term is?  candles PER m**2
output =     cd/m**2
 - a quantity of luminance

Term is?  millicandlepower per are
output =  10 ucd/m**2
 - a quantity of luminance

Term is?  microcandles per hectare
output =  100 pcd/m**2
 - a quantity of luminance

Term is?  megacandela per microhectare
output =  100 Mcd/m**2
 - a quantity of luminance

Term is?  ng/ml
output =   mg/m**3
 - a quantity of density

Term is?  kg/m**3
output =   kg/m**3
 - a quantity of density

Term is?  tonne/micron
output =   Tg/m
 - a quantity of mass per unit length

Term is?  m**3/g
output =  dam**3/kg
 - a quantity of specific volume

Term is?  kA/dm
output =  10 kA/m
 - a quantity of magnetic field strength

Term is?  kiloamps/decametre
output =  100 A/m
 - a quantity of magnetic field strength
```

```
Term is?  hectoamperes per cm
output =  10 kA/m
 - a quantity of magnetic field strength

Term is?  megaA per decimetre
output =  10 MA/m
 - a quantity of magnetic field strength

Term is?  weber per sq m
output =    T
 - a quantity of magnetic flux density, the "tesla"

Term is?  V.s/sq m
output =    T
 - a quantity of magnetic flux density, the "tesla"

Term is?  kilowatthours per gramme
output =  3600 MGy
 - a quantity of absorbed dose, the "gray"

Term is?  MWh/Tg
output =  3.6 Gy
 - a quantity of absorbed dose, the "gray"

Term is?  cycles/s
output =    Hz
 - a quantity of frequency, the "hertz"

Term is?  kilocycles per microsecond
output =   GHz
 - a quantity of frequency, the "hertz"

Term is?  megacycles per usec.
output =   THz
 - a quantity of frequency, the "hertz"

Term is?  gigacycles per picosecond
Wow! You've exceeded SI limits!
Retry.

Term is?  picoC/barn
output =  10 PC/m**2
 - a quantity of electric flux density

Term is?  maxwells per amp
output =  10 nH
 - a quantity of electric inductance, the "henry"
```

Here we see that obsolete units can be used with the SI prefixes, for coherency of thinking, and automatically swallowed up and converted. If a constant preceded the unit, it would be appropriately adjusted.

SLIDES 66 - 70 (ambiguity)

66

```
Term is?  kilofeet/millisecond
output =  304.8 km/s
 - a quantity of velocity

Term is?  fathoms/s
output =  1.8288 m/s
 - a quantity of velocity
```

67

```
Term is?  kilodynes per centimetre
output =    N/m
 - a quantity of surface tension

Term is?  hectopoundals/km
output =  13.8255 mN/m
 - a quantity of surface tension
```

68

```
Term is?  a/v
output =    S
 - a quantity of electrical conductance, the "siemens"

Term is?  gilbert/volt
output =  0.7957747 S
 - a quantity of electrical conductance, the "siemens"
```

69

```
Term is?  square metres per das
output =  10 dam**2/s
 - a quantity of kinematic viscosity

Term is?  sq in/ms
output =  0.64516 m**2/s
 - a quantity of kinematic viscosity
```

70

```
Term is?  t/cubic metre
Enter "t" if you mean "tonne";
"T" if you mean "tesla":  t
output =   Mg/m**3
 - a quantity of density

Term is?  gm/cubic metre
Enter "g" if you mean "gram";
"Gm" if you mean "gigametre2:  g
output =    g/m**3
 - a quantity of density
```

DUALCASE AMBIGUITY

Text processing languages hold the key to easy database usage for handling both customary and SI units and items intermixed. I've developed a method whereby a certain character or symbol is assigned as a line or record divider. Preceding this prefix character are the "types" and "attributes" of the content; following it is the content itself. Thus we have what amounts to a semantic label for every data item in a database.

SLIDE (71) -

Unit Price	Level	System SI	System CU	Stock	Reqd/ Level	Dash No.	Part No.	Name
$45.40	1	X				02	1026	Whatisit
	1		X			02	1026	Whatisit
	2		X	22	1	04	3841	Appurtenance
	3		X	406	6		28732	Bolt
	3		X	2980	6		28666	Nut
	2	X		4	1	02	4864	Appurtenance
	3	X		187	6		28168	Bolt
	3	X		556	6		28170	Nut
	3	X	X	14	2	04	5656	Flange
	2	X		16	1	08	9444	Protuberance
(etc.). . .								

Here we have the columnar form for a parts list and explosion. Only the two rightmost fields constitute "text" information. That is, we could arbitrarily change the part number and name, without affecting anything at all. The other fields are suitable for the prefix information, because any change to them is a real change to content type and values. Note that 1026 is a "use-either" part; that is, it can be used interchangeably after either type of manufacture. So both versions have the same part number. 3841 and 4864 are separate parts; they perform essentially the same function, but are not interchangeable. So they have different part numbers. 5656 is a "don't-care" part, usable to make either the customary unit appurtenance or the SI appurtenance. Here's the prefix form of the same data.

SLIDE (72)-

```
L1SID02\1026-Whatisit
L1CUD02\1026-Whatisit
L2CUD04R1S22\3841-Appurtenance
L3CUR6S406\28732-Bolt
L3CUR6S2980\28666-Nut
L2SID02R1S4\4864-Appurtenance
L3SIR6S197\28168-Bolt
L3SIR6S556\28170-Nut
L3SICUD04R2S14\5656-Flange
L2SID08R1S16\9444-Protuberance
      (etc.). . .
```

The reverse slash is the prefix character, closing off a field that you can see is of variable length. All the text processor cares about is knowing what that character is, and whether you want the part to the left or to the right -- prefix or text. If you specify the prefix, the window opens at the beginning of the line or record, and closes when the prefix character is encountered. If you specify text, the window opens at the prefix character, and closes at the end of the line or record. The text processor cares only whether or not a given string exists in the windowed space. So if I say to subset the file on the prefix "SI", out fall only the parts for metric manufacture. But some of them could be built to customary units! You can see how this approach will permit graceful and leisured conversion without wasting stock.

The columnar form and the prefix form can be mapped to each other in either direction by the text processor, particularly from prefix to columnar for display and records. Unlike the columnar form, however, the prefix form can have any additional descriptive information added to the prefix, in <u>any</u> place within the prefix!

Now let's see how the coherency works in the computational world. Here is a power problem for solution by using both measurement systems.

SLIDE (73) -

PROBLEM

A generator supplies an effective 300 kW (357 kW at 84% efficiency). It's driving engine is 30% efficient.

a) What horsepower is required to drive it? For fuel having calorific value of 18,000 BTU per pound, what's the fuel consumption in pounds per hour?

b) What power is required to drive it? For fuel having calorific value of 42 megajoules per kilogram, what's the fuel consumption in kilograms per hour?

As you see, problem a is in customary units, while problem b is in metric units. For both we need the conversion of 3600 seconds in an hour. For the customary units, we also need:

SLIDE (74) -

1 hp = 746 watts
= 550 Ft-lb/sec
1 BTU = 778 Ft-lb

These are the calculations necessary for customary units.

SLIDE (75) -

WATTS TO KILO

Horsepower required = $\frac{357 \times 1000}{746}$ = 480 hp

Power input at 30% = $\frac{480}{0.3}$ = 1600 hp

Heat content of fuel = $\frac{1600 \times 550}{778}$ BTU/sec

Fuel consumption = $\frac{1600 \times 550}{778 \times 18000}$ lb/sec

= $\frac{1600 \times 550 \times 3600}{778 \times 18000}$ lb/hr

= 227 lb/hr

And here is the same calculation in SI units.

SLIDE 76 -

Power input at 30% = $\frac{357}{0.3}$ = 1190 kW

Heat content of fuel = 1190 kJ/s

Fuel consumption = $\frac{1190 \times 1000 \times 3600}{42 \times 1000000}$ kg/hr

= 102 kg/hr

Note how easy the second calculation is, due to coherency; one watt (power) = one joule (quantity of heat) per second (time). In this case we could substitute and combine into a single expression, where the meaning of each term would be comprehensible.

That absence of conversion factors unclouds the relationships, thus bettering understanding.

Conclusion

You will find the SI a universally powerful tool for unlocking the barriers to understanding. Welcome it and use it, and your productivity will increase. The net costs of conversion are likely to be less than our present fears expect. The computer and careful planning can be combined to make the transition not too painful.

Here's to a Metric America.

SLIDE 77 - logo from AMJ

R Bemer

Yes, It's Inevitable -- and You WILL Enjoy It!

R. W. Bemer, Honeywell Information Systems
Phoenix, AZ 85005, US

We've been forced (economically, by the Common market) to convert to the metric system. SI, that is. With the change come trauma and expectation of heavy expenses. Relax. It's going to be fun when you see how the metric system will make many things so easy, including thinking. And there will be big opportunities for actually saving more than conversion costs.

Now an avid supporter of the SI (Système International), I've changed measurement systems twice, because the SI is not the CGS metric system that I learned first. But I understand people's reluctance to learn a new system when they have the old one so well in hand, and use it so effectively.

For example, the matter of expressing volume. The SI has only the litre and cubic metre, whereas our old system has a profusion. In fact, the historical reason for having several units may be the capability to express volume in units of one. Instead of saying so many litres, we can just say a firkin, a hogshead, a barrel, or one barrel. And everyone knows that a barrel contains 31.5 gallons.

SLIDE 3 - 31.5

Except if it contains petroleum products, in which case it's 42 gallons.

SLIDE 4 - 42

And except for malt beverages, in which case it's 31 gallons.

SLIDE 5 - 31

Of course these barrels are for liquids, but the barrel is also defined for dry content, measured in quarts rather than gallons. Everyone knows that a barrel for dry content contains 105 quarts.

SLIDE 6 - 105

Except if it contains cranberries, in which case it's 87 cuarts.

SLIDE 7 - 87

It can be interesting to see the correspondence, so we convert according to official definition:

SLIDE 8 - conversion

A dry cuart has the same volume as 1.1636 liquid quarts. Calculating 105 times 1.1636,

SLIDE 9

divided by 4 quarts to the gallon, we apparently discover a new barrel, of 30.54 gallons. Let's call it the "barrel for dry liquids"!

SLIDE 10 - "dry liquids"

The public may think that this is a variation in one uncommon unit, and that such inconsistencies do not exist in, say, linear measure. So we say that mathematicians look for the lowest common denominator in linear measure. In doing so, we observe that the old system is again rich in terms, whereas the metric system user has only the metre.

SLIDE 11 - Lowest Common Denominator

1 statute mile	=	4	quarters
	=	8	furlongs
	=	80	chains (survey)
	=	176	chains (football)
	=	320	rods
	=	880	fathoms
	=	1 760	yards
	=	5 280	feet
	=	7 040	spans
	=	8 000	links
	=	15 840	hands
	=	21 120	palms
	=	28 160	nails or ells
	=	63 360	inches
	=	70 400	lines (buttons)
	=	84 480	digits
	=	190 080	barleycorns
	=	760 320	lines
	=	3 041 280	irons
	=	4 055 040	ounces
	=	4 561 920	points
	=	63 360 000	mils

perhaps you think I'm mistaken in showing the number of ounces per mile? Not so. The ounce is equivalent to 1/64 of an inch, and is used in measuring leather thickness!

When one thinks he has found the lowest common denominator for this set, there will always be another to confound it. The "engineer's chain", for example, is either 50 feet or 100 feet. And I haven't even introduced all units of the textile industry -- spindle, hank, heer, cut, skein, lea, bolt, roll, and thread. 220 spindles are 1890 miles long!

But the system is obvicusly rich in relationships. For example:

SLIDE 12

If the average horse is 15.84 hands tall
(and we know that 1 mile = 15 840 hands),
does 1 mile = 1000 horses?

And that's for just the statute mile. How about sea horses and the nautical mile?

Even the foot is subject to variance. Perhaps the Philadelphia foot is the strangest of all.

SLIDE 13 - PHL

It's 1.002 statute feet in the East-West direction, and 1.0045 statute feet North-to-South! This is nontrivial, for it pertains to the city maps that tell you, for example, where the sewer pipes and the telephone cables are buried!

Enough about the old system! You can find many more oddities yourself. But imagine the time wasted in learning these relationships, and especially the computational difficulties that may arise!

The SI Aids Learning

The metric system should help everyone in learning mathematics, engineering, and other quantitative disciplines. The concepts of both number base and exponentiation are built into the system, in the mechanism of prefixes.

SLIDE 14 - Metric Prefixes

10 to +i	i	10 to -i
exa (E)	18	atto (a)
peta (P)	15	femto (f)
tera (T)	12	pico (p)
giga (G)	9	nano (n)
mega (M)	6	micro (&uuu)
kilo (k)	3	milli (m)
hecto (h)	2	centi (c)
deka (da)	1	deci (d)

"Exa" and "peta" may be new to some. They were adopted quite recently, to fill out the range. And what a range it is! Sufficient for all galactic measurement, and all the way down to subatomic measurement.

Observe "nano", which is only 10 to the -9th. Do you know how far light and electricity travel in a nanosecond? Let's calculate it. The speed of light is very close to

SLIDE 15 - 300 Mm/s (megametres/second)

We see "nano-" defined as 10 to the -9, and "mega-" as 10 to the 6.

SLIDE 16 - (prefixes again)

Because 6 - 9 = -3, which we see from the table is "milli-", the speed of light equals

SLIDE 15 - (back up to speed of light)

300 mm/ns (millimetres/nanosecond)

Simple, isn't it? Light travels 300 millimetres in a nanosecond -- just 5 mm short of the old foot.

I chose this example because computers now operate at nanosecond speeds and less, so it's a tough problem to get the results anywhere by wires that must be under 300 mm in length. Despite microcomputers, we still have some computers and peripherals that fill a good-sized room.

Of course, as many people have hand calculators today, we could have done it in the old English system by multiplying 186,283 by 5280, watching the decimal point, then shifting it left 9 places. But it's not as easy as the metric way.

SLIDE 16 - (forward to prefixes again)

Note that above 3 there are no powers except multiples of 3. This practice breeds better comprehension, like marking off three's in writing numbers of many digits. Note also, as a memory convenience, that all symbols are upper case for powers greater than +3.

SLIDE 17 - (distraction)

The SI Abhors Fractions

How about those fractions? The SI doesn't have fractions as such. Do many people have hand calculators that work in fractions? So they can do problems like:

> If Sam can dig 32 ditches in 9 hours, and Joe can dig 24 ditches in 7 hours, how long will it take them to dig one ditch - together? If they can keep out of each other's way?

I have often wondered whether such problems have much real applicability -- if perhaps they just were dreamed up to teach the difficult art of fractions. Fractions may not be difficult to you, but they are to the French, who cannot understand where that 0.375 inch dimension came from, because 3/8ths means nothing at all to them. My company had to import teaching materials from Italy to show the French why such numbers existed, because it is illegal to sell nonmetric materials in France.

The only fraction the French use is 1/2. They will order a demilitre of wine, but a smaller amount is not a quarter litre, which would be 2.5 decilitres, but rather an even 2 decilitres. So you see that our fractional system is the awkward octal or hexadecimal one of computers. Theirs goes 2, 1, 0.5, 0.2, 0.1, 0.05, etc., to form a sequence that repeats in 10's.

And fractions are difficult to my wife, because I have to scale our recipes up and down, and I think difficult to most people. Ask yourself what half of 5 5/8 is. Got it? It's actually a very useful figure.

SLIDE 18 - 2x6

You would use it to saw a 2x6 in half, because the rough size of 6 inches becomes 5 5/8 inches when finished. How much easier to ask "What's half of 140?"

The Why of Converting to the SI

We'll come back to the value of the metric system to the learning and calculating processes, after a digression to find why the United States is converting at just this time. After all, the metric system has been our only legal system for over 100 years, and all of our customary units are defined by their metric values to avoid chaos.

SLIDE 19 - Cost too much?

Legality is apparently not sufficient to avoid a few minor problems. For example, I have my mass as 83 kilos and my height as 185 cm on my driver's license. The last time I was checked by a highway patrolman, he looked at the license and asked how tall I was. I said 185, and he said 185 what? I said centimetres. He said "Come on, now, what's your real height?

So could it be that we have finally realized that the SI, the metric system, is really so much better and easier that we should convert now despite the possibly high cost of that change?

Not in this democracy! The real reason is:

SLIDE 20 -

1978 January 01

The date when non-metric products
may not be sold in Europe!

The US Government may be getting around to it only now, but General Motors, IBM, Ford, Honeywell, Rockwell, and hundreds of other companies have been converting for a long time. The reason is plain. Trade, and the balance of payments factor, which translates directly into jobs and well-being for the American people.

Metric expert Kenyon Taylor, of Beloit Tool, appearing on the TODAY show in New York, said that the loss of export if we didn't convert to metric design would amount to about 5 million jobs!

General Motors is very serious about export.

SLIDE 21 - GM plastic a" card

You might as well get used to the unit "kilopascal", because that's how we'll measure tire pressure. Don't worry about the gas pump reading litres -- the dollar amount is what counts.

Conversion and Dollars

General Motors has one answer to the question "Doesn't it cost too much to change?"

SLIDE 22 - Fanbelts

They now have over 950 different fanbelt sizes that they (and service stations) must stock. When fully converted to metric manufacture, there will be fewer than 50 different sizes! Here's a clear case where a profit is turned on conversion, not a loss. Your garage will be more likely to have your size in stock if the required inventory is smaller. Don't forget, over 40% of the cars on U.S. roads today are metric.

In many other cases such as this, forced redesign has yieldec economic benefits that were lurking there all the time, if we hacn't been so blind. They may often exceed conversion costs. In manufacturing, where most conversion costs were thought to lie, ways have been found to minimize them, and spread them over a period of years.

Conversion and People

The impact upon people is a larger problem. There are many diehards who will fight the SI. The best way to rebut them is to avoid conversion entirely (because it's a crutch), and particularly to avoid exact conversion. Most exact conversion is not justified because of inherent tolerances.

SLIDE 23 - 2" (maybe) nail

Those nails are just as likely to be chopped off at exactly 50 millimetres as they are at exactly 2 inches.

The diehards often use exact conversion to deride the metric system. Somehow football seems to be a favorite target:

SLIDE 24 -

"Bones made the 9.144
metres for first down"

But they don't mind singing

SLIDE 25 -

"I love you --

2219.91 cubic inches and
554.84 cubic inches"

which translates to:

SLIDE 26 -

"I love you --

a bushel and a peck"

Whomever they love that much, that 2774.2 cubic inches of "you" (assuming they just barely float in water) comes to just 100 pounds -- which is OK if their love is 5' 3".

The diehards don't have to worry about football. That is one sport that will never be converted, even though the Canadian football field is 100 metres long. Most U.S. stadiums have the stands crowding the 100 yards and goal areas as closely as possible, and it would be prohibitively expensive to change. Your children, and their children, will still play football, and they will just think of the "yard" as an archaic term peculiar to football, just as you and I think of the "cubit" as an archaic term peculiar to ark-building.

During the course of legislation, conversion was raised as a bugaboo by various union groups. The carpenters, for example, wanted the US to give each of them a huge allowance to buy new tools. This prompted the NBS to make up some displays like this:

SLIDE 27 - Metric and Oldfashioned Hammers

SLIDE 28 - Metric Plane

Of course, a journeyman carpenter starts out with tools worth much less than the allowance they requested.

The electricians couldn't be too recalcitrant about converting to the SI. They've been using it right along!

SLIDE 29 - Metric light bulb

There will be those that feel there are too many new things to learn in the metric system -- particularly the prefixes. Here are some words starting with "P", as taken from the Style Manual of the US Government Printing Office, which is still hopelessly out of cate with respect to the SI.

SLIDE 30 -

Which of these English words
is a unit of length/distance?

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometer	prismometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Can't you just hear the man on the street saying

SLIDE 31 -

"Is a parameter more or less than a meter?"
"How many perimeters in a parameter?"

And according to the formation rules of the SI he would be pronouncing it: "Is a <u>para meter</u> more or less than a meter?" "How many <u>peri meters</u> in a <u>para meter</u>?" After all, the racio announcer doesn't say "ki-LAW-hertz", and the scientists don't say "ki-LAW-grams". So "ki-LAW-meter" is wrong, too.

Wouldn't it be easier this way?

SLIDE 32 -

Is it any easier now?

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometre	prisometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Now do you spot "picometre" (1/100th of an angstrom)?

SLIDE 33 - micrometer

That's pronounced "A MI-cro-ME-tre is so small that it takes a mi-CROM-e-ter to measure it". This difference in pronunciation carries over into hyphenation, so you can tell which is which when reading. Moreover, the international spelling M.E.T.R.E distinguishes between instruments and such, and the unit of length or distance. Handy, isn't it?

The major argument for the "ER" was phonetics. But English is not a phonetic language, and the English spelling is "RE". German is a phonetic language, and it is true that they spell the unit of length "METER". But they don't have our problem! For the instrument, they use the word "Zahler", meaning "counter". In fact, Norway is now converting to the "RE" spelling, even though they have used the metric system for many years.

Can you imagine the assembly line worker at American Motors, putting the 4.0 litre hardware on Gremlins? How will he know to put on the "4.0 litter" for the US, and the "4.0 litre" for export?

The SI is Coherent

The chief pedagogical value of the metric system is its coherency.

SLIDE 34 - Australian Chart (just for directed graph coherency)

SLIDE 35 - Cube for length-volume-mass

A cube one Benson & Hedges cigarette (100 mm) on a side has a capacity of one litre; the water required to fill it has a mass of one kilogram.

While living in Paris, I found coherency in my hat size. In the U.S. it's 7 5/8, a fractional value that I cannot relate to anything at all. I wrapped a tape measure around my head, read 61 centimetres, ordered a size 61 beret at Galeries Lafayette, and -- voila! -- it fit!

The coherency will also be found in the kitchen, where the confusing pounds, quarts, cups, tablespoons, and teaspoons disappear, to be replaced by millilitres. You won't notice the difference between the 4.9 millilitre teaspoon and the 5 millilitre measure that will replace it. I had to memorize that there were 3 teaspoons per tablespoon, and I still get unsure at times. No memorization is needed when the units are 5 ml and 15 ml. In Europe, it is common to combine ingredients with a mass scale. Here is a metric recipe for your inspection, if not cooking:

SLIDE 36 - Metric cupcakes

Let's look at the metre stick, which is usually marked off in centimetres. Imagine it's a dollar in length.

SLIDE 37 - Metre stick

Halfway for a dollar is 50 cents. Halfway for a metre is 50 centimetres. People learn particularly well when you relate things to money. And in this coherent system the same scale can be used to represent temperature (in degrees Celsius).

SLIDE 38 -

Think of it as water freezing at stony broke, boiling at a dollar. Our body temperature is 37 cents, and the thermostat should now be set to 18 cents or less. Again, the value ratio of 18 cents to the dollar is well understood, and so one gets the "feel" of Celsius temperature. Never convert!

Note the little arrow, sitting uncaptioned at 46 degrees. That's the hottest it gets in my home town (Phoenix) in the summertime. On a trip to Stockholm, I took Icelandic Airlines and made the overnight stop there. On the way to Keflavik airport in the bus the next morning, I sat next to a young man, whom I queried as to the temperature. He replied "5 degrees". I remarked that it was 46 when I left home yesterday. ,space SLIDE 39 - Iceland (N pole)

When I assured him that it was indeed Celsius, his eyes became very large. I suppose such a temperature to an Icelander would remind him more of their volcanoes.

The same scale also works for understanding "mass".

SLIDE 40 -

SLIDE 41 - Keyboard (starting computer part)

Computers and the SI

We should see some extensive modification of computer usage in the transition to the SI. The effect upon hardware will be slight, mainly in the output display area. Standard ISO 2955, "Information Processing -- Representations of SI and other units for use in systems with limited character sets", gives replacement symbols, mostly to get around the upper-case-only problem. But most new devices are dual-case now, and at this moment there is a second-draft proposed 8-bit ASCII standard in circulation (X3L2/77/08); in particular, it has the "mu" for micro-, "omega" for ohm, and the degree symbol.

We expect little trouble with Numerical Control. The APT language is already provided with the necessary metric specification capability. The three areas of software impact are conversion and learning, databases, and program rewrite to take advantage of SI coherency to reduce running costs.

The NBS provides a computer program to convert from customary to SI units, but I expect that it will be superseded by programs of the type I shall describe now -- suitable for both conversion and training.

Here are some possibilities in CAL (Computer-Assisted Learning) programs. Imagine dialing your friendly computing service from your terminal.

SLIDES 42 - 48

```
Term is?  x/y
Retry.

Term is?  a/bc
 - a quantity of no meaning in the SI

Term is?  m/s
output =    m/s
 - a quantity of velocity

Term is?  M/MS
output =   um/s
 - a quantity of velocity

Term is?  meter/sec
A "meter" is an instrument.
The unit of length is the "metre".
output =    m/s
 - a quantity of velocity

Term is?  mm per s.
output =   mm/s
 - a quantity of velocity

Term is?  cm/nanosec.
output =  10 Mm/s
 - a quantity of velocity

Term is?  kg.m per s**2
output =    N
 - a quantity of force, the "newton"

Term is?  MN.m/km**3
output =   kJ/m**3
 - a quantity of energy density

Term is?  n.m/s
output =    W
 - a quantity of power, the "watt"

Term is?  newton-metres per seconds
output =    W
 - a quantity of power, the "watt"

Term is?  n.m/k
output =    J/K
 - a quantity of entropy

Term is?  N.M/K
output =    J/K
 - a quantity of entropy
```

Although only the output is shown here, a simple option can be called to also display input mistakes, if any.

SLIDES 49 - 65

```
Term is?  W/M.K
output =     W/(m.K)
 - a quantity of thermal conductivity

Term is?  n.m/KG.k
output =     J/(kg.K)
 - a quantity of specific entropy

Term is?  a.S per v
output =     F
 - a quantity of electrical capacitance, the "farad"

Term is?  v/a
output =     ohm
 - a quantity of electric resistance, the "ohm"

Term is?  vs/amperes
Retry.

Term is?  volts/amps
output =     ohm
 - a quantity of electric resistance, the "ohm"

Term is?  w/a
output =     V
 - a quantity of electromotive force, the "volt"

Term is?  v.s/a
output =     H
 - a quantity of electric inductance, the "henry"

Term is?  V/m
output =     V/m
 - a quantity of electric field strength

Term is?  W/SR
output =     W/sr
 - a quantity of radiant intensity

Term is?  lumens per square metre
output =     lx
 - a quantity of illuminance, the "lux"
```

```
Term is?  cd.sr/square metres
output =     lx
 - a quantity of illuminance, the "lux"

Term is?  candelas/scuare m
output =     cd/m**2
 - a quantity of luminance

Term is?  candlepower per sq m
output =     cd/m**2
 - a quantity of luminance

Term is?  cancles PER m**2
output =     cd/m**2
 - a quantity of luminance

Term is?  millicandlepower per are
output =  10 ucd/m**2
 - a quantity of luminance

Term is?  microcandles per hectare
output =  100 pcd/m**2
 - a quantity of luminance

Term is?  megacandela per microhectare
output =  100 Mcd/m**2
 - a quantity of luminance

Term is?  ng/ml
output =    mg/m**3
 - a quantity of density

Term is?  kg/m**3
output =    kg/m**3
 - a quantity of density

Term is?  tonne/micron
output =    Tg/m
 - a quantity of mass per unit length

Term is?  m**3/g
output =  dam**3/kg
 - a quantity of specific volume

Term is?  kA/dm
output =  10 kA/m
 - a quantity of magnetic field strength

Term is?  kiloamps/decametre
output =  100 A/m
 - a quantity of magnetic field strength
```

```
Term is?  hectoamperes per cm
output =  10 kA/m
 - a quantity of magnetic field strength

Term is?  megaA per decimetre
output =  10 MA/m
 - a quantity of magnetic field strength

Term is?  weber per sq m
output =    T
 - a quantity of magnetic flux density, the "tesla"

Term is?  V.s/sq m
output =    T
 - a quantity of magnetic flux density, the "tesla"

Term is?  kilowatthours per gramme
output =  3600 MGy
 - a quantity of absorbed dose, the "gray"

Term is?  MWh/Tg
output =  3.6 Gy
 - a quantity of absorbed dose, the "gray"

Term is?  cycles/s
output =    Hz
 - a quantity of frequency, the "hertz"

Term is?  kilocycles per microsecond
output =   GHz
 - a quantity of frequency, the "hertz"

Term is?  megacycles per usec.
output =   THz
 - a quantity of frequency, the "hertz"

Term is?  gigacycles per picosecond
Wow! You've exceeded SI limits!
Retry.

Term is?  picoC/barn
output =  10 PC/m**2
 - a quantity of electric flux density

Term is?  maxwells per amp
output =  10 nH
 - a quantity of electric inductance, the "henry"
```

Here we see that obsolete units can be used with the SI prefixes, for coherency of thinking, and automatically swallowed up and converted. If a constant preceded the unit, it would be appropriately adjusted.

SLIDES 66 - 70 (ambiguity)

```
Term is? kilofeet/millisecond
output =  304.8 km/s
 - a quantity of velocity

Term is? fathoms/s
output =  1.8288 m/s
 - a quantity of velocity

Term is? kilodynes per centimetre
output =    N/m
 - a quantity of surface tension

Term is? hectopoundals/km
output =  13.8255 mN/m
 - a quantity of surface tension

Term is? a/v
output =    S
 - a quantity of electrical conductance, the "siemens"

Term is? gilbert/volt
output =  0.7957747 S
 - a quantity of electrical conductance, the "siemens"

Term is? square metres per das
output =  10 cam**2/s
 - a quantity of kinematic viscosity

Term is? sq in/ms
output =  0.64516 m**2/s
 - a quantity of kinematic viscosity

Term is? t/cubic metre
Enter "t" if you mean "tonne";
"T" if you mean "tesla":  t
output =   Mg/m**3
 - a quantity of density

Term is? gm/cubic metre
Enter "g" if you mean "gram";
"Gm" if you mean "gigametre2:  g
output =    g/m**3
 - a quantity of density
```

Text processing languages hold the key to easy database usage for handling both customary and SI units and items intermixed. I've developed a method whereby a certain character or symbol is assigned as a line or record divider. Preceding this prefix character are the "types" and "attributes" of the content; following it is the content itself. Thus we have what amounts to a semantic label for every data item in a database.

SLIDE 71 -

Unit Price	Level	System SI	System CL	Stock	Reqc/ Level	Dash No.	Part No.	Name
$45.40	1	X				02	1026	Whatisit
	1		X			02	1026	Whatisit
	2		X	22	1	04	3841	Appurtenance
	3		X	406	6		28732	Bolt
	3		X	2980	6		28666	Nut
	2	X		4	1	02	4864	Appurtenance
	3	X		187	6		28168	Bolt
	3	X		556	6		28170	Nut
	3	X	X	14	2	04	5656	Flange
	2	X		16	1	08	9444	Protuberance
	(etc.). . .							

Here we have the columnar form for a parts list and explosion. Only the two rightmost fields constitute "text" information. That is, we could arbitrarily change the part number and name, without affecting anything at all. The other fields are suitable for the prefix information, because any change to them is a real change to content type and values. Note that 1026 is a "use-either" part; that is, it can be used interchangeably after either type of manufacture. So both versions have the same part number. 3841 and 4864 are separate parts; they perform essentially the same function, but are not interchangeable. So they have different part numbers. 5656 is a "don't-care" part, usable to make either the customary unit appurtenance or the SI appurtenance. Here's the prefix form of the same data.

SLIDE 72 -

```
L1SID02\1026-Whatisit
L1CUD02\1026-Whatisit
L2CUD04R1S22\3841-Appurtenance
L3CUR6S406\28732-Bolt
L3CUR6S2980\28666-Nut
L2SID02R1S4\4864-Appurtenance
L3SIR6S197\28168-Bolt
L3SIR6S556\28170-Nut
L3SICUD04R2S14\5656-Flange
L2SID08R1S16\9444-Protuberance
      (etc.). . .
```

The reverse slash is the prefix character, closing off a field that you can see is of variable length. All the text processor cares about is knowing what that character is, and whether you want the part to the left or to the right -- prefix or text. If you specify the prefix, the window opens at the beginning of the line or record, and closes when the prefix character is encountered. If you specify text, the window opens at the prefix character, and closes at the end of the line or record. The text processor cares only whether or not a given string exists in the windowed space. So if I say to subset the file on the prefix "SI", out fall only the parts for metric manufacture. But some of them could be built to customary units! You can see how this approach will permit graceful and leisured conversion without wasting stock.

The cclumnar form and the prefix form can be mapped to each other in either direction by the text processor, particularly from prefix to columnar for display and records. Unlike the columnar form, however, the prefix form can have any additional cescriptive information added to the prefix, in <u>any</u> place within the prefix!

Now let's see how the ccherency works in the computational world. Here is a power problem for solution by using both measurement systems.

SLIDE 73 -

PROBLEM

A generator supplies an effective 300 kW (357 kW at 84% efficiency). It's driving engine is 30% efficient.

a) What horsepower is required to drive it? For fuel having calorific value of 18,000 BTU per pound, what's the fuel consumption in pounds per hour?

b) What power is required to drive it? For fuel having calorific value of 42 megajoules per kilogram, what's the fuel consumption in kilograms per hour?

As you see, problem a is in customary units, while problem b is in metric units. For both we need the conversion of 3600 seconds in an hour. For the customary units, we also need:

SLIDE 74 -

1 hp = 746 watts
= 550 Ft-lb/sec
1 BTU = 778 Ft-lb

These are the calculations necessary for customary units.

SLIDE 75 -

Horsepower required = $\frac{357 \times 1000}{746}$ = 480 hp

Power input at 30% = $\frac{480}{0.3}$ = 1600 hp

Heat content of fuel = $\frac{1600 \times 550}{778}$ BTU/sec

Fuel consumption = $\frac{1600 \times 550}{778 \times 18000}$ lb/sec

= $\frac{1600 \times 550 \times 3600}{778 \times 18000}$ lb/hr

= 227 lb/hr

And here is the same calculation in SI units.

SLIDE 76 -

Power input at 30% = $\frac{357}{0.3}$ = 1190 kW

Heat content of fuel = 1190 kJ/s

Fuel consumption = $\frac{1190 \times 1000 \times 3600}{42 \times 1000000}$ kg/hr

= 102 kg/hr

Note how easy the second calculation is, due to coherency; one watt (power) = one joule (quantity of heat) per second (time). In this case we could substitute and combine into a single expression, where the meaning of each term would be comprehensible.

That absence of conversion factors unclouds the relationships, thus bettering understanding.

Conclusion

You will find the SI a universally powerful tool for unlocking the barriers to understanding. Welcome it and use it, and your productivity will increase. The net costs of conversion are likely to be less than our present fears expect. The computer and careful planning can be combined to make the transition not too painful.

Here's to a Metric America.

SLIDE 77 - logo from AMJ

NATIONAL COUNCIL OF
Teachers of Mathematics
PHOENIX MEETING / 24-26 FEBRUARY 1977

March 20, 1977

R. W. Bemer
Honeywell Information Systems
P. O. Box 6000
Phoenix, Arizona 85005

Dear Mr. Bemer,

On behalf of the Board of Directors of the National Council of Teachers of Mathematics and the program committee for the Phoenix name-of-site meeting, thank you for contributing to our successful program. We realize the amount of expertise, effort and talent required to design and present a major address, particularly a presentation which contributes to the personal as well as the professional growth of those in attendance. Again, THANKS!

Warmest regards,

Gene Gardenhire

Gene Gardenhire
for the program committee

CONVENTION COCHAIRMEN

SHIRLEY FRYE, 6802 E. Sunnyvale, Scottsdale, Arizona 85253
AL FERENCZ, Saguaro High School, 6250 N. 82d Street, Scottsdale, Arizona 85253

NCTM MEETINGS COMMITTEE REPRESENTATIVE

GAIL LOWE, 1471 Calle Colina, Thousand Oaks, California 91360

PROGRAM COMMITTEE

GARY BITTER, Chairman, College of Education, Arizona State University, Tempe, Arizona 85281

JERALD MIKESELL
Mesa Public Schools
14 W. Second Avenue
Mesa, Arizona 85201

CARLTON LEE EVANS
102 Pike Avenue
Alamosa, Colorado 81101

GENE GARDENHIRE
3516 Singapore Circle N.E.
Albuquerque, New Mexico 87105

ROSEMARY ANDERSON
University of Nevada
Department of Professional Studies
College of Education
Las Vegas, Nevada 89109

AMMON L. LOUDERMILK
1708 W. Royal Palm Road
Phoenix, Arizona 85021

ESTHER MILNE
Tucson District #1
P.O. Box 4040
Tucson, Arizona 85717

PHOENIX COMMITTEE CHAIRMEN

REGISTRATION

LINDA VOLLSTEDT
Alhambra High School
3839 Camelback Road
Phoenix, Arizona 85019

ROBERT GOODRICH
Alhambra High School
3839 Camelback Road
Phoenix, Arizona 85019

HOSPITALITY

PAUL SHOECRAFT
Department of Mathematics
Arizona State University
Temple, Arizona 85281

SUPPLIES, EQUIPMENT, & MEETING ROOMS

GEORGE PAVOL
Glendale Union High School District
7650 N. 43d Avenue
Glendale, Arizona 85301

PUBLICITY

PHYLLIS STEINMANN
Scottsdale Community College
9000 E. Chaparral Road
Scottsdale, Arizona 85252

FREE MATERIALS

PHIL EVANSTOCK
Trevor Browne High School
7402 W. Catalina Drive
Phoenix, Arizona 85033

CONVENTION NEWSLETTER

TOM HOPKINS
Saguaro High School
6250 N. 82d Street
Scottsdale, Arizona 85253

NCTM MATERIALS

ROWENA SCHATTENBERG
Hohokam Elementary School
8451 East Oak Street
Scottsdale, Arizona 85257

RECEPTION & MEALS

JAN MILLER
Chaparral High School
6935 East Gold Dust Avenue
Scottsdale, Arizona 85253

FILMS & FILMSTRIPS

JACKSON FITZ-RANDOLPH
Madison Number 1
5525 N. 16th Street
Phoenix, Arizona 85016

MEMBERSHIP

RUSSELL JACOBS
Trevor Browne High School
7402 W. Catalina Drive
Phoenix, Arizona 85033

EVALUATION

STEPHEN PACE
Safford High School
Box 960
Safford, Arizona 85546

FINAL REPORT

BILL WILLIAMS
Grand Canyon College
3300 W. Camelback Road
Phoenix, Arizona 85019

SCHOOL EXHIBITS

JOHN RUCKER
Central High School
4525 N. Central Avenue
Phoenix, Arizona 85012

STUDENT AIDES

RON MCCULLY
Phoenix Union High School District
2526 W. Osborn Road
Phoenix, Arizona 85017

SIGNS & PRINTING

SCOTT BULL
Camelback High School
4612 N. 26th Street
Phoenix, Arizona 85016

NATIONAL COUNCIL OF
Teachers of Mathematics
1906 Association Drive, Reston, Virginia 22091 ■ (703) 620-9840

10 May 1978

Mr. R. W. Bemer
Honeywell Information Systems
P. O. Box 6000
Phoenix, AZ 85005

Dear Mr. Bemer:

Thank you for your letter of 9 May giving us the approval to proceed with processing your article "Metrication Aids Education--and Vice Versa." As noted on the enclosed letter of 1 September, I believe that you have in your possession an edited copy that was sent to you. The thrust of the reviewers' comments was that the manuscript should be shortened in some way, and the marked manuscript was a compilation of suggestions for doing this. If you do not have this copy, please let me know and I will try to reconstruct the comments and send them to you. You did send the necessary artwork for the article in your letter of 25 April 1977.

Also enclosed is a copyright release form which we would like you to sign and return. If you have any questions about it, please let me know.

I hope to hear from you at your convenience.

Sincerely yours,

Harry B. Tunis

Harry B. Tunis
Managing Editor
The MATHEMATICS TEACHER

HBT:bbc
Enclosures

NATIONAL COUNCIL OF
Teachers of Mathematics
1906 Association Drive, Reston, Virginia 22091 ■ (703) 620-9840

19 April 1978

Mr. R. W. Bemer
Honeywell Information Systems
P. O. Box 6000
Phoenix, AZ 85005

Dear Mr. Bemer:

The Board of Directors of the National Council of Teachers of Mathematics has adopted a policy to permit the publishing of material using the liter or litre spellings. Please let us know if this change will permit you to have us proceed with the publication of your material "Metrication Aids Education--and Vice Versa."

Thank you for your interest in the MATHEMATICS TEACHER.

Sincerely yours,

Harry B. Tunis
Managing Editor
The MATHEMATICS TEACHER

HBT:bbc

NATIONAL COUNCIL OF
Teachers of Mathematics
1906 Association Drive, Reston, Virginia 22091 ■ (703) 620-9840

4 October 1977

Mr. R. W. Bemer
Honeywell Information Systems
P.O. Box 6000
Phoenix, AZ 85005

Dear Mr. Bemer:

The Editorial Panel of the MATHEMATICS TEACHER has considered your request to use -re spellings for meter in your article "Metrication Aids Education--and Vice Versa." Recent action by the Board of Directors requires Council publications to follow the -er spelling. We hope that you will consider the suggestions in our letter of 1 September and rethink your position on the spelling of meter so that we may process your material for publication.

Thank you very much for your patience during these extended deliberations.

Sincerely yours,

Harry B. Tunis
Managing Editor
The MATHEMATICS TEACHER

HBT:drv

NATIONAL COUNCIL OF
Teachers of Mathematics
1906 Association Drive, Reston, Virginia 22091 ■ (703) 620-9840

19 September 1977

Mr. RW Bemer
Honeywell Information Systems
P. O. Box 6000
Phoenix, AZ 85005

Dear Mr. Bemer:

I am sorry that the suggested changes in our letter of 1 September 1977 concerning your article "Metrication Aids Education--and Vice Versa" were not acceptable. Unless I hear from you to the contrary, I would like to bring the spelling issue to the full Panel at the upcoming meeting on 1 October. Perhaps some proposal can be reached on how to proceed. The points in your paper are certainly well made and would be of great interest to our readers.

I hope this upcoming reconsideration is agreeable.

Sincerely yours,

Harry B. Tunis

Harry B. Tunis
Managing Editor
The MATHEMATICS TEACHER

HBT:drv

Honeywell

1977 September 13

Mr. Harry B. Tunis,
Managing Editor
The MATHEMATICS TEACHER
1906 Association Drive
Reston, Virginia 22091

Sir:

I have received your acceptance of my paper "Metrication Aids Education", conditional upon using the spelling "er" to reflect Federal Register comments on the SI.

Unknown to you, of course, I have a substantial international reputation in the computer business. I was the United States representative in the compilation of the International Vocabulary for Information Processing (IFIP). I am a personal friend of the Secretary General of ISO, where custody of the SI lies.

My answer to you, sir, is that as an author I should be permitted to use words according to my own conscience. I reject your offer to prostitute (Webster 3rd, Meaning 2.).

I have published over 75 papers so far, and do not need to blemish my professional position. Representatives of the U. S. Government have been in error before, as they are now in this matter. But I do not intend to follow them blindly.

So consign your copy of my paper to the wastebasket, as I have mine. The matter is closed.

RBemer

RW Bemer

cc: Dr Gary Bitter, ASU
Dr Jordan Baruch, Asst. Secy of Commerce,
Science and Technology
Shirley Frye, Scottsdale Schools
Louis Sokol, Pres., US Metric Assn
Olle Sturen, Sec. Gen., ISO

HONEYWELL INFORMATION SYSTEMS, P. O. BOX 6000, PHOENIX, ARIZONA 85005, TELEPHONE 602/993-6000

State University of New York at Buffalo

553 CHRISTOPHER BALDY HALL, AMHERST, NEW YORK 14260 TEL.(716)636-2451

DEPARTMENT OF INSTRUCTION

FACULTY OF EDUCATIONAL STUDIES

February 15, 1978

Mr. Louis F. Sokol
U.S. Metric Association, Inc.
Sugarloaf Star Route
Boulder, Colorado 80302

Dear Mr. Sokol:

Re your letter to John Egsgard of January 11: I agree. The action to support -er spelling only was in my opinion one of the more foolish actions of NCTM. It happened before any current board member was serving. Because of that action (and, more important, because metric activity has claimed an inordinate and unwarranted amount of time and energy of the mathematics education community) I have not participated in board discussions of or actions on metrics during my tenure of office.

In response to your reasonable concern, however, I am prepared to entertain this subject at the San Diego board meeting. I will enter the following motions:

MOVED That the National Council of Teachers of Mathematics rescinds its earlier policy specifying a single spelling of the two metric units, meter and liter, and their derivatives. The result of this action is to make NCTM policy toward this matter of orthography neutral, that is, that beyond matters of consistency it is up to individual authors to use the spelling they prefer for

meter or metre

and

liter or litre

and their derivatives (kilometer or kilometre, for example).

MOVED That the new permissive policy on metric spelling will be implemented immediately but only as it affects future publications, including journals, and revisions of current publications when major expense is not involved. It will be communicated to NCTM membership in an early issue of the NEWSLETTER, MATHEMATICS TEACHER, ARITHMETIC TEACHER, JOURNAL FOR RESEARCH IN MATHEMATICS EDUCATION and BULLETIN FOR LEADERS with an accompanying brief rationale including but not limited to recognition of -re spelling in Canada and by the U.S. Metric Association.

I cannot predict the response to these motions, of course. If you wish to provide me with information that I could use in support of the main motions, please feel free to do so. I must note, however, that I am prepared neither to make nor to support a motion to fix the spelling on metre and litre or even to indicate a preference for this usage. I prefer to leave lit- and met- just like theat-, open to -er or -re.

GRR/dlt
CC: John Egsgard — President
Shirley Hill — President-elect
James Gates — Executive Director

Sincerely yours,

Gerald R. Rising

Gerald R. Rising
Professor

Dear Bob:

I wonder if there is some way for NCTM members to support Prof Rising's motion? I am sure he is embarrassed with the NCTM policy.

Regards,
Lou

I talked to Valene about your NCTM letter, and she felt we should not publish it at this time.

Honeywell Interoffice Correspondence

Date: 5 May 1977

To: R. W. Bemer - Phoenix

From: Gayle Miels - Corp. Research Center (Jim Lufkin)

Location: MN09

Subject: "METRICATION AIDS EDUCATION"

Bob,

I'd walk a hundred furlongs in my new 2-ounce sandals to hear you lecture math teachers on metrication.

I'm sure Jim has no objection to your publishing this delightful piece in Mathematics Teacher magazine (he's out of town this week).

It's the best thing that came across my desk today.

Saturday

7:30 a.m.-8:45 a.m. BREAKFAST Coronado

Presider: Gary Bitter, Arizona State University, Tempe, Arizona

Speaker: R.W. BEMER, Honeywell Information Systems, Phoenix, Arizona

Metrication Helps Education -- and Vice Versa

Educators wish to impart conceptual understanding, not rote memorization that passes a course but won't serve one's work in life. Incoherency in OUR/THE old measuring system was a barrier to insight and understanding...a turnoff to most. Metrication should be welcomed, enjoyed, and espoused with every enthusiasm, for seldom is there such an opportunity for a measurable increase in teaching productivity.

CONTINUITY COPY

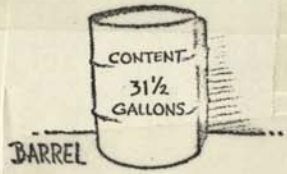
METRICATION AIDS EDUCATION -- AND VICE VERSA

R. W. Bemer, Honeywell Information Systems
Phoenix, AZ 85005, US

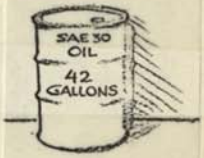
(adapted from a lecture at the 1977 February 24-26 meeting, National Council of Teachers of Mathematics.)

Now an avid supporter of the SI (Système International), I've changed measurement systems twice, because the SI is not the CGS metric system that I learned first. But I understand people's reluctance to learn a new system when they have the old one so well in hand, and use it so effectively.

For example, the matter of expressing volume. The SI has only the litre unit, whereas our old system has several. In fact, the historical reason for having several units might be the capability to express volume in units of one. Instead of saying so many litres, we can just say a firkin, a hogshead, a barrel, or one barrel. And everyone knows that a barrel contains 31.5 gallons.



Except if it contains petroleum products, in which case it's 42 gallons.



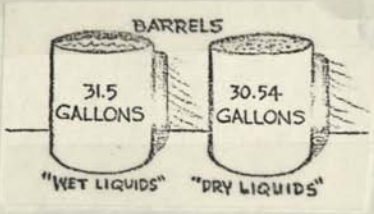
And except for malt beverages, in which case it's 31 gallons.

Of course these barrels are for liquids, but the barrel is also defined for dry content, measured in quarts rather than gallons. Everyone knows that a barrel for dry content contains 105 quarts.

Except if it contains cranberries, in which case it's 87 quarts.

It can be interesting to see the correspondence, so we convert according to official definition -- a dry quart has the same volume as 1.1636 liquid quarts.

Calculating 105 times 1.1636, divided by 4 quarts to the gallon, we apparently discover a new barrel, of 30.54 gallons. Let's call it the "barrel for dry liquids"!



You may think that this is a variation in one uncommon unit, and that such inconsistencies do not exist in, say, linear measure. Let us take the mathematician's viewpoint, because we are such, and seek the lowest common denominator in linear measure. In doing so, we observe (Table 1) that the old system is again rich in terms, whereas the metric system user has only the metre.

1 statute mile	=	8 furlongs
	=	80 chains (survey)
	=	176 chains (football)
	=	320 rods
	=	880 fathoms
	=	1 760 yards
	=	5 280 feet
	=	7 040 spans
	=	8 000 links
	=	15 840 hands
	=	63 360 inches
	=	84 480 digits
	=	4 055 040 ounces
	=	63 360 000 mils

Table 1. Mathematicians Seek The Lowest Common Denominator

You think I'm mistaken in showing the number of ounces per mile? Not so. The ounce is equivalent to 1/64 of an inch, and is used in measuring leather thickness!

When the student thinks he has found the lowest common denominator for this set, there will always be another to confound it. The "engineer's chain", for example, is either 50 feet or 100 feet. But the system is obviously rich in relationships. For example:

If the average horse
is 15.84 hands tall,
and 1 mile = 15 840 hands,

Does 1 mile = 1000 horses?

And that's for just the statute mile. How about sea horses and the nautical mile?

One "knot" =
47 (statute) feet, 3 inches

The number of rope knots
paid out in 28 seconds =

nautical miles per hour

Again, as mathematicians, we verify the relationship:

$$\frac{\text{3600 seconds}}{\text{28 seconds}} = \frac{\text{1 nautical mile}}{\text{47' 3"}}$$

Thus a nautical mile = 6075'

But,
American nautical mile = 6076.11549'
British nautical mile = 6080'

So we're not too exact in our knot-tying! There are other variations. The Roman mile is 1620 yards, but don't rush to say that this equals 4860 feet, because the Roman pace, which equals 5 Roman feet, is equal to only 4.85 English feet. Imagine how surprised the Romans might have been when they conquered Londinium and its bigger-footed inhabitants!

We see that even the foot is subject to variance. Perhaps the Philadelphia foot is the strangest of all, being 1.002 statute feet in the East-West direction, and 1.0045 statute feet North-to-South! This is nontrivial, for it pertains to the city maps that tell you, for example, where the sewer pipes and telephone cables are buried!

That's enough about the old system; you can find many more oddities yourself. But imagine the time wasted in learning these relationships, and especially the computational difficulties that may arise!

The SI Aids Learning

Will the metric system help students to learn mathematics and other sciences? I believe that it will. The concepts of both number base and exponentiation are built into the system, in the mechanism of prefixes (Table 2).

10^{+i}	i	10^{-i}
exa (E)	18	atto (a)
peta (P)	15	femto (f)
tera (T)	12	pico (p)
giga (G)	9	nano (n)
mega (M)	6	micro (μ)
kilo (k)	3	milli (m)
hecto (h)	2	centi (c)
deka (da)	1	deci (d)

Table 2. Metric Prefixes

"Exa" and "peta" may be new to some. They were adopted quite recently, to fill out the range. And what a range it is! Sufficient for all galactic measurement, and all the way down to subatomic measurement.

Observe "nano", which is only 10 to the -9th. Do you know how far light and electricity travel in a nanosecond? Let's calculate it. The speed of light is very close to

300 Mm/s (megametres/second)

Table 2 defines "nano-" as 10 to the -9, and "mega-" as 10 to the 6. Because 6 - 9 = -3, which we see from the table is "milli-", the speed of light equals

300 mm/ns (millimetres/nanosecond)

Simple, isn't it? Light travels 300 millimetres in a nanosecond -- just 5 mm short of the old foot.

If you wonder why this example was chosen, remember that computers now operate at nanosecond speeds and less, so it's a tough problem to get the results anywhere by wires that must be under 300 mm in length. Despite microcomputers, we still have some computers and peripherals that fill a good-sized room.

Of course, inasmuch as all students have hand calculators today, we could have done it in the old English system by multiplying 186,283 by 5280, watching the decimal point, and then shifting it left 9 places. But it's not as easy as the metric way.

Note (also in Table 2) that above 3 there are no powers except multiples of 3. This practice breeds better comprehension, like marking off three's in writing numbers of many digits. Note also, as a memory convenience, that all symbols are upper case for powers greater than +3.

The SI Abhors Fractions

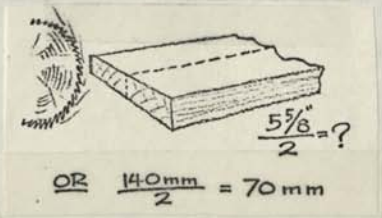
How about those fractions? Do all students have hand calculators that work in fractions? So they can do problems like:

> If Sam can dig 32 ditches in 9 hours, and Joe can dig 24 ditches in 7 hours, how long will it take them to dig one ditch - together? If they can keep out of each other's way?

I have often wondered whether such problems have much real applicability -- if perhaps they just were dreamed up to teach the difficult art of fractions. Fractions may not be difficult to you, but they are to the French, who cannot understand where that 0.375 inch dimension came from, because 3/8ths means nothing to them. My company had to import teaching materials from Italy to show the French why such numbers existed, because it is illegal to sell nonmetric materials in France.

The only fraction the French use is one half. They will order a demilitre of wine, but a smaller amount is not a quarter litre, which would be 2.5 decilitres, but rather an even 2 decilitres. So you see that our fractional system is the awkward octal or hexadecimal one of computers. Theirs goes 2, 1, 0.5, 0.2, 0.1, 0.05, etc., to form a sequence that repeats in 10's.

And fractions are difficult to my wife, because I have to scale our recipes up and down, and I think difficult to most people. Ask yourself what half of 5 5/8 is. It's actually a very useful figure. You would use it to saw a 2x6 in half, because the rough size of 6 inches becomes 5 5/8 inches when finished. How much easier to ask "What's half of 140?"



The Why of Converting to the SI

We'll come back again to the value of the metric system to the learning processes, after a digression to question why the United States is converting at just this time. After all, the metric system has been our only legal system for over 100 years, and all of our customary units are defined by their metric values to avoid chaos.

Legality is apparently not sufficient to avoid a few minor problems. For example, I have my mass as 85 kilos and my height as 185 cm on my driver's license. The last time I was checked by a highway patrolman, he looked at the license and asked how tall I was. I said 185, and he said 185 what? I said centimetres. He said "Come on, now, what's your real height?

So could it be that we have finally realized that the SI, the metric system, is really so much better and easier that we should convert now despite the possibly high cost of that change?

Not in this democracy! The real reason is:

1978 January 01

The date when non-metric products may not be sold in Europe!

The US Government may be getting around to it only now, but General Motors, IBM, Ford, Honeywell, Rockwell, and hundreds of other companies have been converting for a long time. The reason is plain. Trade, and the balance of payments factor, which translates directly into jobs and well-being for the American people.

Kenyon Taylor of the U.S. Metric Conversion Board, appearing on the TODAY show in New York, said that the loss of export if we didn't convert to metric design would amount to about 5 million jobs!

Figure 1 is an example of how serious General Motors is about export.

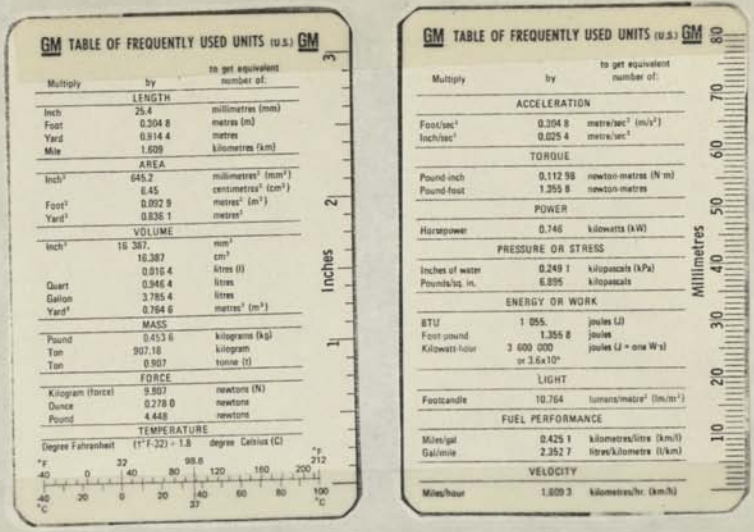


Figure 1. General Motors is Ready

You might as well get used to the unit "kilopascals", because that's how we'll measure the pressure in your tires. Don't worry about the gas pump reading litres -- the dollar amount is the thing that counts.

As teachers, you may be proud to know that the state school systems formed the second major sector to appreciate the need for conversion to the SI, via the 33-state Consortium on Metric Education. I've talked to hundreds of teachers this last year. And I've seen enough confusion, variation, and mistakes in teaching materials to convince me that nationwide coordination is essential. In this regard I recommend support for the only organization with this goal -- NAME -- the National Association for Metric Education. Lola May, its president, is well known to the NCTM.

Conversion and Dollars

General Motors has one answer to the question "Doesn't it cost too much to change?" They now have over 950 different fan belt sizes that they (and service stations) must stock. When fully converted to metric manufacture, there will be fewer than 50 different sizes! Here's a clear case where a profit is turned on conversion, not a loss. Your garage will be more likely to have your size in stock if the required inventory is smaller. Don't forget, over 40% of the cars on U.S. roads today are metric.

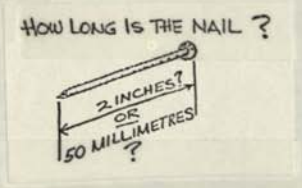
In many other cases such as this, forced redesign has yielded economic benefits that were lurking there all the time, if we hadn't been so blind. They may often exceed conversion costs. In manufacturing, where most conversion costs were thought to lie, ways have been found to minimize them, and spread them over a period of years.

As metric manufacture begins, one can compute the corresponding settings on scales calibrated in customary units. Later these scales can have overlays applied. Then new metric machine tools can be purchased without added expense, because such tools do wear out, and become obsolete via the newer numerically controlled machines. Gradually, as the percentage of metric manufacture increases, this process can be repeated.

Conversion and People

The impact upon people is more critical. There are many diehards who will fight the SI. The best way to rebut them is to avoid conversion entirely (because it's a crutch), and particularly to avoid exact conversion.

Most exact conversion is not justified because of inherent tolerances. Those nails are just as likely to be chopped off at exactly 50 millimetres as they are at exactly 2 inches.



The diehards often use exact conversion to deride the metric system. Somehow football seems to be a favorite target. Opponents of the metric system will sneer:

> "Bones made the 9.144 metres for first down"

But they don't mind singing

> I love you --
>
> 2219.91 cubic inches and
> 554.84 cubic inches

Which translates to:

I love you --

a bushel and a peck

Whomever they love that much, that 2774.2 cubic inches of "you" (assuming they just barely float in water) comes to just 100 pounds -- which is OK if their love is 5' 3".

The diehards don't have to worry about football. That is one sport that will never be converted, even though the Canadian football field is 100 metres long. All U.S. stadiums have the stands crowding the 100 yards and goal areas as closely as possible, and it would be prohibitively expensive to change. Your children, and their children, will still play football, and they will just think of the "yard" as an archaic term peculiar to football, just as you and I think of the "cubit" as an archaic term peculiar to ark-building.

During the course of legislation, conversion was also raised as a bugaboo by various union groups. The carpenters, for example, wanted the US to give each of them a huge allowance to buy new tools (see Figures 2 and 3).

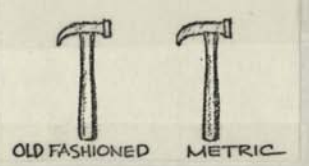


Figure 2. Types of Hammers

Figure 3. Metric Plane

Of course, a journeyman carpenter starts out with tools worth much less than the allowance they requested.

Figure 4 shows why the electricians couldn't be too recalcitrant about converting to the SI. They've been using it right along!

Figure 4. Converted Light Bulb

There will be those that feel there are too many new things to learn in the metric system -- particularly the prefixes. Table 3 shows some words starting with "P", as taken from the Style Manual of the US Government Printing Office, which is still hopelessly out of date with respect to the SI.

Which of these English words
is a unit of length/distance?

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometer	prisometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Table 3. P.....METERS

Can't you just hear the man on the street saying

Is a parameter more or less than a meter?

How many perimeters in a parameter?

And according to the formation rules of the SI he would be pronouncing it this way: "Is a para meter more or less than a meter?" "How many peri meters in a para meter?"

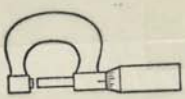
Wouldn't it be easier as shown in Table 4?

Is it any easier now?

pantometer	phacometer	polymeter
parameter	phorometer	porometer
passimeter	phytometer	potometer
passometer	picometre	prisometer
pathometer	piezometer	pulmometer
pedometer	pitometer	pulsimeter
peirameter	planimeter	pulsometer
pelvimeter	planometer	pycnometer
pentameter	pleximeter	pyrometer
perimeter		

Table 4. Correct SI Spelling

Now do you spot the "picometre", which is 1/100th of an Angstrom?



a micrometRE is so small ...
you need a micrometER to measure it

That's pronounced "A MI-cro-ME-tre is so small that it takes a mi-CROM-e-ter to measure it". This difference in pronunciation carries over into hyphenation, so you can tell which is which when reading. Moreover, the international spelling M.E.T.R.E distinguishes between instruments and such, and the unit of length or distance. Handy, isn't it? So why did the NCTM prefer the M.E.T.E.R spelling?

The NCTM, represented by
Joseph R. Caravella
cast 3 votes for the spelling

METER

The major argument for the "ER" was phonetics. But English is not a phonetic language, and the English spelling is "RE". German is a phonetic language, and it is true that they spell the unit of length "METER". But they don't have our problem! For the instrument, they use the word "Zahler", meaning "counter". In fact, Norway is now converting to the "RE" spelling, even though they have used the metric system for many years.

Can you imagine the assembly line worker at American Motors, putting the 4.0 litre hardware on Gremlins? How will he know to put on the "4.0 litter" for the US, and the "4.0 litre" for export?

We need all the aids we can get in making the conversion to the metric system. Perhaps the NCTM can change its collective mind on this matter.

The SI is Coherent

The chief pedagogical value of the metric system is its coherency, as demonstrated in Figures 5 and 6. A cube one Benson & Hedges cigarette (100 mm) on a side has a capacity of one litre; the water required to fill it has a mass of one kilogram.

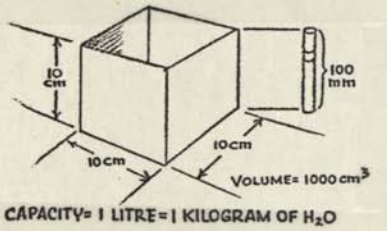


Figure 5. Volume-Mass Coherency

LENGTH

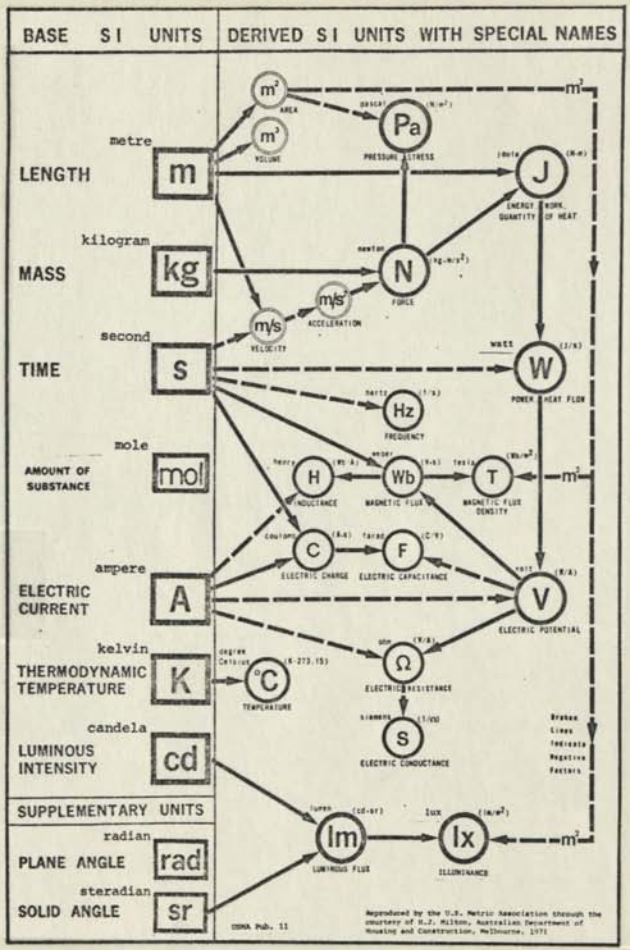


Figure 6. Systemwide Coherency - The Australian Chart

While living in Paris, I found coherency in my hat size. In the U.S. it's 7 5/8, a fractional value that I cannot relate to anything at all. I wrapped a tape measure around my head, read 61 centimetres, ordered a size 61 beret at Galeries Lafayette, and -- voila! -- it fit!

The coherency will also be found in the kitchen, where the confusing pounds, quarts, cups, tablespoons, and teaspoons disappear, to be replaced by millilitres. You won't notice the difference between the 4.9 millilitre teaspoon and the 5 millilitre measure that will replace it. I had to memorize that there were 3 teaspoons per tablespoon, and I still get unsure at times. No memorization is needed when the units are 5 ml and 15 ml. In Europe, it is common to combine ingredients with a mass scale. Figure 7 is a metric recipe for your inspection, if not cooking:

HAPPY DAY CUPCAKES

625 mℓ sifted cake flour
375 mℓ sugar
15 mℓ baking powder
5 mℓ salt
125 mℓ butter (room temperature)
250 mℓ milk
5 mℓ vanilla
2 eggs

Measure sifted flour, add sugar, salt, and baking powder. Place shortening in mixing bowl; stir to soften. Sift in dry ingredients. Add 175 mℓ milk, beat 2 minutes on medium speed (by hand beat 150 strokes per minute). Add eggs, vanilla, and remainder of milk. Beat 1 more minute. Bake cupcakes for 20 minutes at 175 °C. Cool, then frost with favorite frosting.

Figure 7. Metric Recipe

Let's look at the metre stick (Figure 8), which is usually marked off in centimetres. Imagine it's a dollar in length. Halfway for a dollar is 50 cents. Halfway for a metre is 50 centimetres.

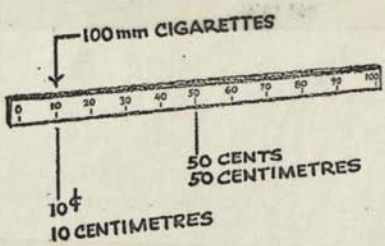


Figure 8. Metre Stick

I've noticed that people learn particularly well when you relate things to money. And in this coherent system the same scale can be used to represent temperature (in degrees Celsius). Figure 9 shows water freezing at 0, boiling at a dollar. Our body temperature is 37 cents, and the thermostat should now be set to 18 cents or less.

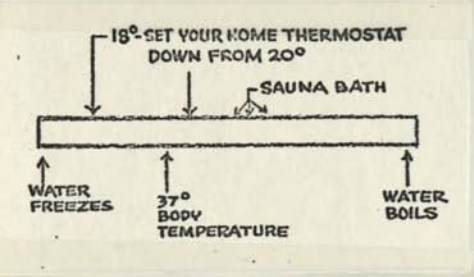


Figure 9. Relative Temperature

Again, the value ratio of 18 cents to the dollar is well understood, and so one gets the "feel" of Celsius temperature. Never convert!

Note the little arrow, sitting uncaptioned at 46 degrees. That's the hottest it gets in my home town (Phoenix) in the summertime. On a trip to Stockholm, I took Icelandic Airlines and made the overnight stop there. On the way to Keflavik airport in the bus the next morning, I sat next to a young man, whom I queried as to the temperature. He replied "5 degrees". When I remarked that it was 46 when I left home yesterday, he said "Of course you mean Fahrenheit". When I assured him that it was indeed Celsius, his eyes became very large. I suppose such a temperature to an Icelander would remind him more of their volcanoes.

Figure 10 shows that the same scale also works for understanding "mass".

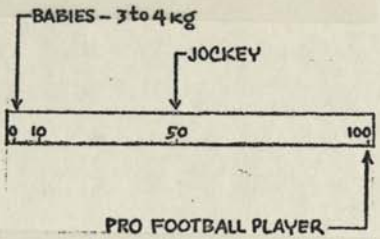


Figure 10. Relative Mass

Now let's see how the coherency works in the computational world [1]. Figure 11 poses a problem for solution by using both measurement systems.

PROBLEM

A generator supplies an effective 300 kW (357 kW at 84% efficiency). It's driving engine is 30% efficient.

a) What *horsepower* is required to drive it? For fuel having calorific value of 18,000 *BTU per pound,* what's the fuel consumption in *pounds* per hour?

b) What *power* is required to drive it? For fuel having calorific value of 42 *megajoules per kilogram,* what's the fuel consumption in *kilograms* per hour?

Figure 11. A Power Problem

As you see, problem a is in customary units, while problem b is in metric units. For both we need the conversion of 3600 seconds in an hour. For the customary units, we also need:

1 hp = 746 watts
= 550 Ft-lb/sec

1 BTU = 778 Ft-lb

Figure 12 shows the calculations necessary for customary units.

$$\text{Horsepower required} = \frac{357 \times 1000}{746} = 480 \text{ hp}$$

$$\text{Power input at 30\%} = \frac{480}{0.3} = 1600 \text{ hp}$$

$$\text{Heat content of fuel} = \frac{1600 \times 550}{778} \text{ BTU/sec}$$

$$\text{Fuel consumption} = \frac{1600 \times 550}{778 \times 18000} \text{ lb/sec}$$

$$= \frac{1600 \times 550 \times 3600}{778 \times 18000} \text{ lb/hr}$$

$$= 227 \text{ lb/hr}$$

Figure 12. Calculation in Customary Units

Figure 13 shows the same calculation in metric units.

$$\text{Power input at 30\%} = \frac{357}{0.3} = 1190 \text{ kW}$$

$$\text{Heat content of fuel} = 1190 \text{ kJ/s}$$

$$\text{Fuel consumption} = \frac{1190 \times 1000 \times 3600}{42 \times 1000000} \text{ kg/hr}$$

$$= 102 \text{ kg/hr}$$

Figure 13. Calculation in Metric Units

Note how easy the second calculation is, due to coherency; one watt (power) = one joule (quantity of heat) per second (time). In this case we could substitute and combine into a single expression, where the meaning of each term would be comprehensible.

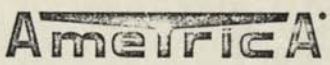
That absence of conversion factors unclouds the relationships, thus bettering the student's understanding. A vivid experience from my early life occurred when college vacation preceded high school vacation by a week, and my father (the school superintendent) asked me to teach a math course instead of the regular teacher. I found them hung up on binomial expansion, and asked "Didn't anyone ever show you Pascal's triangle?". In 30 minutes they knew what they did not learn in the whole previous week!

Conclusion

That's the way it is when one uses the metric system -- the SI. You will find it a universally powerful tool for unlocking the barriers to understanding. Welcome it and use it, and your teaching productivity will increase. To put it another way, you may not need to work so hard, and the remaining work will be more pleasurable as your students are not turned off by computational complexities.

Here's to a Metric America.

A



REFERENCE

1. J.L.Pokorney, "Metrication and systems design", Honeywell Computer Journal 7, No. 2, 1973, 123-126.