MSPM SECTION BZ.10.04

Identification
APL Character Conversion

Purpose
The APL character set is far removed from ANSCII. Although it is Multics policy to remain close to the ANSCII standard, the de facto popularity of APL with its own established character set indicated that Multics APL should where necessary depart from ANSCII in order to be as compatible as possible with other APL implementations. Since Multics device interface modules do not know of the special APL character set, the necessary conversions are done by the APL interpreter.

Terminals
APL supports the same two basic terminals as does Multics: the Model 37 Teletype-writer (and other ANSCII devices indistinguishable from it, like the GE Terminet 300), and the IBM 1050 (also 2741) with PTTC/BCD keyboard.

IBM 1050 terminals are assumed by APL to be equipped with the APL typing element (IBM part number 1167988). The normal APL character set is thus available to users of such a terminal. Multics, of course, assumes that the terminal is equipped with EBCDIC graphics (ball 963), and translates the input to ANSCII on that basis. APL, however, knowing the relative positions of characters on the 963 and 988 balls, translates again to produce APL internal codes.

ANSCII terminals are presumed not to have interchangable graphics. Users of such terminals must use an ANSCII representation of the APL character set which is described below. This representation is, to be sure, more inconvenient to use than an IBM 1050 with the proper printing element, but it does allow usage of APL from a far greater variety of terminals.

Editing Characters
Erase, kill, and escape characters are provided by the APL character converter. The normal editing of the standard Multics device interface module is turned off, because the standard Multics editing characters fall in inconvenient places in the APL character set.

The kill character is 100-octal, represented as alpha on 1050s and @ on Model 37s. The erase character is 043-octal, represented as omega on 1050s and # on Model 37s. The escape character is 042-octal, represented as diaeresis on 1050s and " on Model 37s.

A great many escape sequences are defined for Model 37s. This is an attempt to provide a mnemonic means of typing the APL character set on ANSCII devices. In each case, the escape sequence consists of the first two letters of some name of the character or its function (two first letters if two words). Note also that the 26 alphabetic characters underscored in the APL character set are represented as upper-case alphabetics on Model 37s. Lower-case alphabetics represent the 26 plain alphabetics.

Other Escapes

In addition to the escape sequences defined in the device tables, four other escape sequences are also accepted. They are:

"@	represents @ in the input line with no kill of previous characters;
"#	represents # in the input line with no erasure of previous characters;
""	represents " in the input line with no escape of next character; and
"xyz	where x, y, and z are octal digits, represents the octal code xyz in the input line.

Special Functions

The APL character converter performs two special functions not performed at all by the normal Multics device interface modules. One is performed on input, and one on output.

First, on input, the character converter can be given a set of characters and a carriage position which represent the state of the current line before the user is given an opportunity to type some more on it. Then the characters typed by the user will be merged with the set originally on the line to form the effective output line. This feature is used to support the change-character editing feature of APL editing.

Second, on output, the character converter can be given an integer, which is the index of some character of the output line. When the converter converts the line to a physical stream of characters, the print position finally occupied by that character will be returned to the caller (it may occupy a print position less than its original index because of previous backspace characters, or more because of previous characters which must be expanded into escape sequences). This feature is used to support the typing of a mark under the offending character of a line in error.

Device Tables

The following table lists the APL internal code and meaning down the left, followed by the way the character must be typed on an IBM 1050 (or 2741) equipped with a 988 ball, followed by the way the character must be typed on a Model 37 Teletype (or other ANSCII device).

APL	1050	37
010 bs	bs	bs
011 ht	ht	ht
012 nl	nl	nl
040 sp	sp	sp
041 !	! .	!
042 ¨	¨	"
043 ⍵	⍵	#
044 ⍳	⍳	$
045 ⍴	⍴	%
046 ×	×	&
047 '	'	'
050 (	(	(
051)	)	)
052 *	*	*
053 +	+	+
054 ,	,	,
055 -	-	-
056 .	.	.
057 /	/	/
060 0	0	0
061 1	1	1
...		
071 9	9	9
072 :	:	:
073 ;	;	;
074 <	<	<
075 =	=	=
076 >	>	>
077 ?	?	?

APL	1050	37
100 ⍺	⍺	@
101 A	_ A	A
102 B	_ B	B
...		
132 Z	_ Z	Z
133 [	[	[
134 \	\	\
135]	]	]
136 ¯		^
137 _	_	_
140 ∇	∇	`
141 *A*	*A*	a
142 *B*	*B*	b
...		
172 *Z*	*Z*	z
173 ←	←	{
174 |	|	|
175 →	→	}
176 ~	~	~
200 ≤	≤	< _
201 ≥	≥	> _
202 ≠	≠	/ =
203 ∨	∨	"or
204 ∧	∧	"an
205 ÷	÷	- :
206 ∈	∈	"ep
207 ↑	↑	"up

APL	1050	37
210 ↓	↓	"do
211 ○	○	"ci
212 ⌈	⌈	"ce
213 ⌊	⌊	"fl
214 Δ	Δ	"de
215 ⍟	⍟	"cc
216 ⍞	⍞	"qu
217 ∩	∩	"ca
220 ⊥	⊥	"ev
221 ⊤	⊤	"en
222 ⊂	⊂	"in
223 ⊃	⊃	"co
224 ∪	∪	"cu
225 ⍱	~ ∨	"no
226 ⍲	~ ∧	"na
227 ⍝	- ○	"rf
230 ⌿	- /	- /
231 ⍒	∇ ~	"lf
232 ⍟	* ○	"lo
233 ⌽	| ○	"rr
234 ⍉	\ ○	"tr
235 ⍬	*O T U* o t u	
236 ⍋	∇ |	"gd
237 ⍙	| Δ	"gu
240 ⍲	⍟ ∩	"la
241 ⍞	' ⍞	"qq
242 ⌶	⊥ ⊤	"ib
243 ⍀	- \	- \

Communication overstriken