

問題2 次の文字データとその表現に関する記述を読み、各設問に答えよ。

コンピュータでは、特定のビットパターンを割り当てることで文字を表現する。半角の英数字や記号は1バイトを用い、漢字やひらがな等の全角文字の場合は2バイト以上を用いる。このように表現されたものを文字コードと呼ぶ。文字コードには、利用したい文字の集まりである文字集合と、どの文字集合に合わせて変換するかを決定するエンコードの概念が含まれている。

<設問1> 次の1バイト文字データに関する記述中の に入れるべき適切な字句を解答群から選べ。

現在使われている1バイトの文字コードはASCIIコードが世界的に利用されている。これは、ANSI（米国標準規格協会）で制定された7ビットを基にしており、8ビット目はデータ通信に用いるときのパリティビットとしている。

8ビットで表現できる文字コードは256種類だが、その半分はメーカーや各国の標準化団体が独自の文字を割り当てている。日本では、ASCIIコードを基にし、半角のカタカナを収録したJIS X 0201（ANKコード）が使われている。

次の図は、ASCIIコードに収録されている文字とコードを対応させたものである。なお、 $(000\ 0000)_2 \sim (001\ 1111)_2$ （16進数の00～1F）、および $(111\ 1111)_2$ （16進数の7F）は制御文字である。

								0	0	0	0	1	1	1	1
								0	0	1	1	0	0	1	1
								0	1	0	1	0	1	0	1
b7	b6	b5	b4	b3	b2	b1	列 行	0	1	2	3	4	5	6	7
			0	0	0	0	0	NUL	DLE	空白	0	@	P	`	p
			0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
			0	0	1	0	2	STX	DC2	"	2	B	R	b	r
			0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
			0	1	0	0	4	EOT	DC4	\$	4	D	T	d	t
			0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
			0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
			0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
		1	0	0	0		8	BS	CAN	(8	H	X	h	x
		1	0	0	1		9	HT	EM)	9	I	Y	i	y
		1	0	1	0		10	LF	SUB	*	:	J	Z	j	z
		1	0	1	1		11	VT	ESC	+	;	K	[k	{
		1	1	0	0		12	FF	FS	,	<	L	¥	l	
		1	1	0	1		13	CR	GS	-	=	M]	m	}
		1	1	1	0		14	SO	RS	.	>	N	^	n	~
		1	1	1	1		15	SI	US	/	?	O	_	o	DEL

図 ASCIIコード

この図から、「Z」が割り当てられているコードは $(101\ 1010)_2$ (16進数の5A) であることがわかる。同様に, 「A」, 「a」, 「3」, 「#」という文字は, それぞれ16進数で次のようなコードになる。

A	…	(1)
a	…	(2)
3	…	(3)
#	…	(4)

また, コードが16進数で順に57, 33, 43となる文字列は(5)である。

(1) ～ (4) の解答群

- | | | | |
|-------|-------|-------|-------|
| ア. 14 | イ. 16 | ウ. 23 | エ. 32 |
| オ. 33 | カ. 41 | キ. 61 | |

(5) の解答群

- | | | | |
|--------|--------|--------|--------|
| ア. u34 | イ. u3C | ウ. W34 | エ. W3C |
|--------|--------|--------|--------|

<設問 2> 次の日本語の文字コードに関する記述に関連の深い字句を解答群から選べ。

(6) Windows などのパソコン上で標準的に用いられる 2 バイトコードである。1 バイト文字と 2 バイト文字の区別は先頭の 1 バイトで判断できるようになっており, シフトイン, シフトアウトのような制御文字を使用しない。

(7) AT&T が日本語 UNIX システム諮問委員会の提案に基づいて, 文字コードの各国定義部分に日本語の文字集合を割り当てたもので, UNIX 上で日本語を利用するため用いられる。

(8) 米国の IT 関連企業が中心となって提唱して標準化された文字コードで, 1 つの文字コード体系で各国の言語を処理するものである。全ての文字を 2 バイトで表現するため半角文字や全角文字を統一的に利用できる。現在は様々な文字が追加され, 一部では 3 バイト以上で定義されている。

(6) ～ (8) の解答群

- | | | |
|--------------|-------------|------------|
| ア. EBCDIC | イ. EUC-JP | ウ. JIS |
| エ. Shift JIS | オ. TRON コード | カ. Unicode |