

# H24. 秋. ネット7-7

## 問3 電子メールで用いる MIME 形式（ネットワーク）

(H24 秋・FE 午後問3)

### 【解答例】

- 〔設問1〕 aーエ, bーエ  
〔設問2〕 オ, カ, ケ

### 【解説】

電子メールの MIME 形式がテーマの問題であるが、基本的な文字符号の変換の問題であるともいえる。電子メールや MIME の知識が少なくても、問題文中の説明を丁寧に読んでいけば、解答できる問題である。

Windows や Mac OS, Linux などの多くの OS で使用されている文字符号（文字コード）は Unicode という形式である。Unicode では、世界中のコンピュータで使用される多言語の文字を収容するために、複数の文字符号を使用できるが、一般的に UTF-8 (UCS Transformation Format-8) が用いられることが多い。UCS (Universal multi-octet coded Character Set) は、いわゆる Unicode のことであるが、その符号化方式には UTF-8, UTF-16, UTF-32 がある。UTF-8 は、8 ビット（バイト）単位で情報を表す 8 ビット符号で、1 バイト～4 バイトの可変長（規格の上では 6 バイトまで）である。しかし、インターネット電子メールは、US-ASCII (US-American Standard Code for Information Interchange) のような 7 ビット符号で書かれたテキストしか送信できない。そのため、UTF-8 などの 8 ビット符号は、MIME (Multipurpose Internet Mail Extensions；マイム) を使用して 7 ビット符号に変換（エンコード）する必要がある。MIME は、多目的にインターネットメールを拡張するための仕様で、7 ビット符号以外のテキストやバイナリ（16 進数）データ、音声や画像データなどを 7 ビット符号に変換（エンコード）する。

なお、文字符号には、“0”～“9”や“A”～“Z”などの表示や印刷が可能な図形文字と、STX（テキストの開始）や CR（改行）などの制御用の制御文字がある。

### 〔設問1〕

- ・空欄 a：MIME の変換（エンコード）方式である、base64 と quoted-printable に関する問題である。問題文にも記述があるように、base64 は、コンテンツを先頭から 6 ビットごとに区切り、各 6 ビットを、変換表に基づいて図形文字（“A”～“Z”、“a”～“z”、“0”～“9”、“+”、“/”）1 文字に変換する。  
変換表は、問題文には示されていないが、参考のために、表 A に示す。なお、図形文字の“=”は、パディング（長さ調整用の詰め物）として使用される。  
quoted-printable は、コンテンツをバイト列とみなし、US-ASCII の制御文字又は図形文字の“=”と一致するバイト、及び先頭ビットが 1 のバイトを、“=XX”（XX は 2 桁の 16 進数数列）の形の 3 文字の US-ASCII 図形文字列に置き換える（一部例外があるが、この問題では考慮しない）。

表 A 変換表

ビット列	文字	ビット列	文字	ビット列	文字	ビット列	文字
000000	A	010000	Q	100000	g	110000	w
000001	B	010001	R	100001	h	110001	x
000010	C	010010	S	100010	i	110010	y
000011	D	010011	T	100011	j	110011	z
000100	E	010100	U	100100	k	110100	0
000101	F	010101	V	100101	l	110101	1
000110	G	010110	W	100110	m	110110	2
000111	H	010111	X	100111	n	110111	3
001000	I	011000	Y	101000	o	111000	4
001001	J	011001	Z	101001	p	111001	5
001010	K	011010	a	101010	q	111010	6
001011	L	011011	b	101011	r	111011	7
001100	M	011100	c	101100	s	111100	8
001101	N	011101	d	101101	t	111101	9
001110	O	011110	e	101110	u	111110	+
001111	P	011111	f	101111	v	111111	/

空欄 a1～a4 に入れる答えは、a に関する解答群の中から組合せとして正しいものを選ぶものとする。そこで、まずは、空欄 a1, a2 について考える。

“@△IPA△2012.”を base64 又は quoted-printable で変換した結果が、“wqkgSVBBIDIwMTIu”又は“=C2=A9△IPA△2012.”になる。

ここで、“=C2=A9△IPA△2012.”に注目すると、“=”から始まっているため、“@=C2=A9”は@を quoted-printable で変換した結果であるとすぐに推測できる。また、残りの“△IPA△2012.”は US-ASCII の図形文字と一致するので、置換えは発生しない。よって、“@△IPA△2012.”を quoted-printable で変換した結果が、“=C2=A9△IPA△2012.”となり、空欄 a1 には quoted-printable が当てはまる。したがって、空欄 a1 は quoted-printable となり、その結果、必然的に空欄 a2 は base64 となる。

念のため、“@△IPA△2012.”を base64 で変換することを考える。

“@”は 16 進数で C2A9 の 2 バイトの符号であるから、C2A9 を 2 進数で表現すると 1100 0010 1010 1001 となる。“△”は空白で、16 進数で 20 であるから、0010 0000 である。

よって、“@△”の 16 進数の表現は 11000010 10101001 00100000 となり、これを 6 ビットごとに区切ると、110000 101010 100100 100000 となる。変換表に当てはめてみると、“wqkg”となる。続く“IPA”は 16 進数で 49, 50, 41 となり、これを 2 進数で表現すると 01001001 01010000 01000001 になる。これを 6 ビットごとに区切ると、010010 010101 000001

000001 となり、変換表から、“SVBB”となる。以下、同様に変換を行うことによって“wqkgSVBBIDIwMTIu”の文字列が得られる。

次に、空欄 a3, a4 について考える。

“@△IPA△2012.”を base64 で変換した結果が“wqkgSVBBIDIwMTIu”（16 文字）となり、quoted-printable で変換した結果が“=C2=A9△IPA△2012.”（16 文字）となる。図形文字である“IPA△2012.”の部分（9 文字）に注目すると、base64 で変換した結果が“SVBBIDIwMTIu”（12 文字）となり、quoted-printable で変換した結果が“IPA△2012.”（9 文字）となる。よって、図形文字の場合は、quoted-printable で変換した方が短くなり、空欄 a3 には quoted-printable が当てはまる。逆に、バイナリデータの場合、quoted-printable で変換すると、1 バイトのデータが“=”から始まる 3 文字に変換されるので長くなる。base64 で変換した方が短くなるので、空欄 a4 には base64 が当てはまる。以上のことから、空欄 a1 には quoted-printable、空欄 a2 には base64、空欄 a3 には quoted-printable、空欄 a4 には base64 が入る。したがって、(エ)が正解である。

- ・空欄 b：UTF-8 では、平仮名 1 文字に先頭ビットが 1 であるバイト三つから成る符号を割り当てている（平仮名 1 文字で 3 バイトとなる）。UTF-8 で書かれた 1 文字の平仮名を quoted-printable で変換すると、先頭ビットが 1 のバイトは“=XX”の形の 3 文字の US-ASCII 図形文字列に置き換えるため、3 バイト×3 文字=9 バイトとなる。UTF-8 で書かれた 6 文字の平仮名を quoted-printable で変換すると、6 文字×9 バイト=54 バイトとなる。したがって、(エ)が正解である。

### 〔設問2〕

MIME 形式を使用すると、図 1 のように複数のコンテンツを格納することができる（これを multipart という）。複数のコンテンツを格納する場合、コンテンツの境目（boundary）を示すための文字列である **delimiter** を指定する必要がある。設問 2 の問題文にもあるように、**--delimiter** がコンテンツの開始、**--delimiter--** がコンテンツの終了を示す。**delimiter** は任意の文字列が使用できるが、コンテンツの内容と **delimiter** の文字列が誤認識されないような文字列を選択する必要がある。

図 2 で示された 3 行をコンテンツに含むとき、**delimiter** としてふさわしくない文字列を解答群の選択肢から選ぶ。

各選択肢を **delimiter** として当てはめた場合における、コンテンツの開始（**--delimiter**）と、コンテンツの終了（**--delimiter--**）を表 B に整理する。

表 B 選択肢の文字列と delimiter

選択肢	文字列	開始 ( <b>--delimiter</b> )	終了 ( <b>--delimiter--</b> )	備考
ア	<b>--case1</b>	<b>----case1</b>	<b>----case1--</b>	
イ	<b>--case2</b>	<b>----case2</b>	<b>----case2--</b>	
ウ	<b>--case3</b>	<b>----case3</b>	<b>----case3--</b>	
エ	<b>case1</b>	<b>--case1</b>	<b>--case1--</b>	
オ	<b>case2</b>	<b>--case2</b>	<b>--case2--</b>	図 2 の <b>--case2</b> が含まれる。
カ	<b>case3</b>	<b>--case3</b>	<b>--case3--</b>	図 2 の <b>--case3--</b> が含まれる。
キ	<b>case1--</b>	<b>--case1--</b>	<b>--case1----</b>	
ク	<b>case2--</b>	<b>--case2--</b>	<b>--case2----</b>	
ケ	<b>case3--</b>	<b>--case3--</b>	<b>--case3----</b>	図 2 の <b>--case3--</b> が含まれる。

以上のことから、(オ)、(カ)、(ケ)が図 2 に示される“コンテンツ 1”に含まれる文字列とぶつかるため、ふさわしくない。したがって、(オ)、(カ)、(ケ)が正解である。