

## トレース初級問 2 ハフマン符号化(データ圧縮) (2004 年春FE午後問 3 改)

問 データ圧縮の方法の一つであるハフマン符号化に関する次の記述を読んで、設問 1, 2 に答えよ。

ハフマン符号化は、圧縮対象の文字列を構成する文字の種類に注目し、その出現回数の偏りを利用した圧縮方法である。今、図 1 に示す 100 文字からなる文字列の圧縮を考える。全角 1 文字を 16 ビット (2 バイト) で表現する方式の場合、この文字列は、100 文字なので、1,600 ビットの長さである。ここで、この文字列を構成する文字の種類とその出現回数は、図 2 に示すとおりであるとする。

あんたがたどこさ△ひごさ△ひごどこさ△ ~ (中略) ~  
くってさ△それをこのはでちょっとかくせ

図1 対象となる文字列(△は空白を示す)

文字の種類	△	さ	て	っ	く	こ	が	せ	た	と	ど	ま	ん
出現回数	12	11	6	5	4	4	3	3	3	3	3	3	3

  

う	ご	そ	で	に	は	ば	ひ	も	や	よ	れ
2	2	2	2	2	2	2	2	2	2	2	2

  

を	あ	い	お	か	き	し	ち	ぬ	の	ぼ	り
2	1	1	1	1	1	1	1	1	1	1	1

図2 文字の種類と出現回数

設問 1 次の記述中の  に入れる正しい答えを記述せよ。

対象となる文字列は、図 2 に示すように、37 種類の文字から構成されている。37 種類の文字を固定長のビット列を用いて識別するためには、1 種類の文字に  ビットのビット列を対応させればよく、対象となる文字列は、 ビット/文字× 100 文字で計算される長さで表現できる。このとき、対象となる文字列の長さは、表に示すように、全角 1 文字を 16 ビットで表現する方式の場合を 100 % とすると、 % となる。

表 文字の表現方式と対象文字列の長さ

表現の方式	1 文字の長さ(ビット)	対象の文字列の長さ(ビット)	圧縮の割合(%)
全角 1 文字を 16 ビットで表現	16	1,600	100
37 種類の文字を固定長で表現	<span style="border: 1px solid black; display: inline-block; width: 60px; height: 15px; vertical-align: middle;"></span> a	(省略)	<span style="border: 1px solid black; display: inline-block; width: 60px; height: 15px; vertical-align: middle;"></span> b



文字の種類	△	さ	て	っ	く	こ	が
出現回数	12	11	6	5	4	4	3
対応ビット列	100	010	1010	0010	0000	11101	10110
ビット長	3	3	4	4	4	5	5
出現ビット長(注)	36	33	24	20	16	15	15
出現ビット累計	36	69	93	113	129	149	164

せ	た	と	ど	ま	ん	う	ご	そ	で
3	3	3	3	3	3	2	2	2	2
00111	00110	01101	01100	01111	01110	00010	101111	101110	110101
5	5	5	5	5	5	5	6	6	6
15	15	15	15	15	15	10	12	12	12
179	194	209	224	239	254	264	276	288	300

に	は	ば	ひ	も	や	よ	れ	を	あ
2	2	2	2	2	2	2	2	2	1
110100	110111	110110	110001	110000	110011	110010	111101	111100	000110
6	6	6	6	6	6	6	6	6	6
12	12	12	12	12	12	12	12	12	6
312	324	336	348	360	372	384	396	408	414

い	お	か	き	し	ち	ぬ	の	ぼ	り
1	1	1	1	1	1	1	1	1	1
0001111	0001110	1110001	1110000	1110011	1110010	1111101	1111100	1111111	1111110
7	7	7	7	7	7	7	7	7	7
7	7	7	7	7	7	7	7	7	7
421	428	435	442	449	456	463	470	477	484

図4 文字の出現回数と対応するビット列に関する情報

## トレース初級問2 ハフマン符号化(データ圧縮) 解答用紙

( )月( )日 時間 ( )分( )秒 目安 15分 点数 ( )点

氏名 \_\_\_\_\_

### 設問1

a. \_\_\_\_\_

b. \_\_\_\_\_

### 設問2

c. \_\_\_\_\_

d. \_\_\_\_\_

e. \_\_\_\_\_