

必須問題

問題 1 次のランレングス符号化に関する記述中の に入れるべき適切な字句または数値を解答群から選べ。

ランレングス符号化とはデータ圧縮に用いられる符号化の一種で、連続する同じ値と、その個数の組み合わせで次のような形式で表現する。

[データ] [個数] [データ] [個数] …

例：X 5 Y 9 … データ X が 5 個連続し、データ Y が 9 個連続する。

ここで扱うデータは 1 バイトの英字とし、連続するデータが 255 個を超える場合は、255 個で区切り符号化する。これにより、データと個数をそれぞれ 1 バイトで表現できる。

なお、問題の構成上、個数は 10 進数で表記する。

また、ここでは圧縮率を次のように定義する。

圧縮率 = 符号化後のデータ量 ÷ 符号化前のデータ量

ただし、圧縮率は小数第 3 位を四捨五入する。

文字列「AAAAAABBBBAAAAAAAAAA」(20 バイト) を符号化すると、文字列は A が 6 個、B が 4 個、A が 10 個なので、符号化すると次のようになる。

A 6 B 4 A 10

符号化前のデータ量は 20 バイトであり、符号化後は 6 バイトであるから、圧縮率は (1) となる。

また、文字列「AABBCDDEFGGH」(12 バイト) を符号化すると、

A 2 B 2 C 1 D 2 E 1 F 1 G 2 H 1

となる。このときの圧縮率は (2) となり、圧縮ではなく、逆に冗長になる。

次に、以下に示す①～⑤の 12 バイトの文字列を符号化したときについて考える。

① AAABCCDDDD

② AABCCDDEEFF

③ AAAAAABCDEEE

④ ABCDDDEEFFG

⑤ AAAABCDEFghi

①～⑤の文字列のうち、もとのデータ量より圧縮できる（圧縮率が 1 未満である）文字列は (3) である。この中で最も圧縮効果が高いのは文字列 (4) であり、圧縮率は 0.67 である。

一方、圧縮が有効でなく冗長となる（圧縮率が 1 を超える）文字列は (5) である。この中で最も冗長性が高くなるのは文字列 (6) であり、圧縮率は 1.50 である。

(1) , (2) の解答群

ア. 0.30 イ. 0.60 ウ. 0.75 エ. 1.33 オ. 3.33

(3) , (5) の解答群

ア. ①と② イ. ①と③ ウ. ①と④ エ. ①と⑤ オ. ②と③
 カ. ②と④ キ. ②と⑤ ク. ③と④ ケ. ③と⑤ コ. ④と⑤

(4) , (6) の解答群

ア. ① イ. ② ウ. ③ エ. ④ オ. ⑤