

(ア) 文字の符号化と文字の出現確率と平均符号長の関係について説明している次の文章の空欄 (23) から (30) (31) に入るもっとも適した数字を解答欄にマークしなさい。

A,B,C,D,E の 5 種類の文字から構成される文字列がある場合に、それを 0 と 1 からなる数字の列に変換して、データとして保存したり通信を行うことを考える。

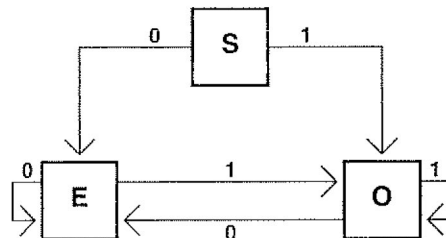
A に 00、B に 01、C に 10、D に 110、E に 111 を割り当てる。

このように割り当てを決めると、ABCAE という文字列は 00011000111 に符号化される。5 文字が 11 ビットに変換されているので、1 文字あたり 2.20 ビットで符号化されていることになる。より長い文字列について考えると、元の文字列中の A,B,C,D,E の出現確率が同じであれば、符号化したときの、1 文字あたりの平均ビット長は、(23) . (24) (25) ビットとなる (小数点以下第 3 位を四捨五入)。

A,B,C,D,E の出現比率が、A:B:C:D:E=1:2:2:2:3 だった場合には、前述の符号化による 1 文字あたりの平均ビット長は、(26) . (27) (28) ビット (小数点以下第 3 位を四捨五入) となり、(23) . (24) (25) ビットより長くなってしまう。これは、出現確率の高い文字 (E) に、長い符号 (3 ビット) が割り当てられているのが原因であり、各文字の出現確率が与えられた場合、平均符号長を短くするには、出現確率の高い文字に短い符号を割り当てるほうがよいことがわかる。

次に、A,B,C,D,E,F の 6 種類の文字から構成される文字列について考える。A,B,C,D,E,F の出現比率が、A:B:C:D:E:F=1:2:2:2:3:5 だった場合、新たに最適な符号化を行うと、1 文字あたりの平均ビット長は、(29) . (30) (31) ビット (小数点以下第 3 位を四捨五入) となる。

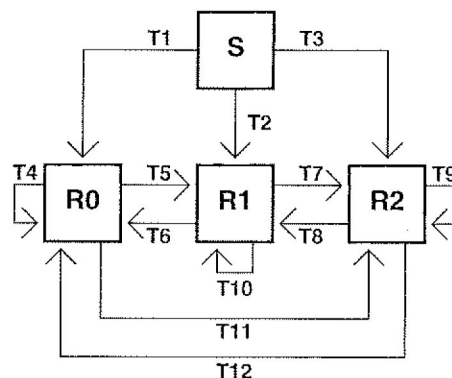
(イ) 任意長の 0 と 1 の列が 1 桁ずつ入力されるものとする。1 桁ずつ入力されるときに先に入力されたものを、2 進法表現の上位の桁として扱い、そこまでに入力された数が、偶数か奇数かを判定して、状態を変更することにする。この入力と状態の変化を示したのが次の図である。



S は開始状態を表し、矢印が状態の変化を示し、その状態の変化を引き起こす入力が、その矢印についている数字 (0 または 1) である。E は入力されたそれまでの 0 と 1 の列が偶数である状態、O は奇数である状態を示す。

たとえば、10110 と入力されると、O → E → O → O → E と、状態が変化していく。

これと同様に、入力された2進法で表現された数を3で割った余りを判定して、3つの状態間を変化させる。余り0、1、2に対して、それぞれ、状態 R0、状態 R1、状態 R2 を割り当てる。



図中の状態の変化 T1 から T12 までを引き起こす入力の値を示すと、次の表になる。解答欄 (32) から (44) には、状態の変化を引き起こす入力が0または1のどちらかの場合は、それぞれ0または1を、相当する状態の変化が起こす入力がない場合には、9を記入しなさい。

状態変化	入力の値	状態変化	入力の値
T1	(32)	T7	(38)
T2	(33)	T8	(39)
T3	(34)	T9	(40)
T4	(35)	T10	(41)
T5	(36)	T11	(42)
T6	(37)	T12	(43)

この図を用いると、桁数の多い2進法で表された数に対しても、それを3で割った余りが、直接計算せずに、状態の変化をたどるだけでわかる。24ビットの2進法表現の数 1110 1110 1110 1110 1110 1110 を3で割った余りを解答欄 (44) に記入しなさい。