

非参照非相談 QuizL06

樋口さぶろお

龍谷大学工学部数理情報学科

情報処理の基礎 L06(2014-10-29 Wed)



<http://hig3.net>

1

Quiz(文字の個数と文字コードのビット長)

以前は、電報に使える文字の種類は、カタカナ 46 個 (小文字の ユヨや長音ーを含まない)・濁点・半濁点・空白・句点・読点の 51 個程度だった。これを最低限の長さのビットパターンに符号化するとすると、何ビット必要か、理由をつけて答えよう。

2

Quiz(遺伝子のビット長)

遺伝子は、A,G,C,T(アデニン, グアニン, シトシン, チミン) の 4 種類の「文字」(本当は塩基)の文字列(本当は DNA の塩基配列)で記述されている。大腸菌の遺伝子は、長さが 4600000 文字(塩基)程度の文字列(塩基配列)である。大腸菌の遺伝子をビットパターンに書き直すとなると、何ビット程度必要か、理由をつけて答えよう。

1

Quiz 略解+コメント:文字の個数と文字コードのビット長

略解

$2^5 < 51 \leq 2^6$ より 6 bit. $\log_2 51$ を繰り上げて整数値にしても同じ答.

コメント

$51_{(10)} = 110011_{(2)}$ で右辺が長さ 6 だから 6bit, という考え方は誤り.

もし文字の個数が 64 個だったら, $64_{(10)} = 1000000_{(2)}$ で長さ 7 だけど, $000000_{(2)}, \dots, 111111_{(2)}$ の 64 個に符号化できるでしょ.

この問題も, $000000_{(2)}, \dots, 110010_{(2)}$ の 51 個に符号化できる.

2

Quiz 略解+コメント:遺伝子のビット長

略解

$4600000 \log_2 4 = 9200000$ bit.