Project Euler 59. XOR Decryption

hiragn

2024 年 12 月 25 日

1. 問題の概要

文字コードの説明

ASCII (American Standard Code for Information Interchange) は文字コードであり，大文字 $A = 65$，アスタリスク $* = 42$，小文字 $k = 107$ のように各文字にコードが割り当てられている。

XOR 暗号の説明

暗号化の方法としてテキストファイルの各文字を ASCII に変換し，秘密鍵から計算された値と XOR（演算子としての記号は $\oplus$）を取る手法がある。暗号化に用いたのと同じ暗号化鍵でもう一度 XOR を取ると平文に戻せる。

$$65 \oplus 42 = 107, \; 107 \oplus 42 = 65$$

問題の説明

暗号化された ASCII のコードを含むファイル `0059_cipher.txt`[a] が与えられる。暗号化鍵は 3 文字の小文字である。
平文はよく用いられる英単語を含んでいる。この暗号文を復号し，平文の ASCII での値の和を求めよ。

https://projecteuler.net/problem=59

[a] [https://projecteuler.net/project/resources/0059_cipher.txt:title

2. 解法

現代英語の最頻単語は”the”らしいです。

- `http://user.keio.ac.jp/~rhotta/hellog/2010-03-01-1.html`
- `http://jbauman.com/gsl.html`

「アルファベット小文字 3 文字の暗号化キーを全部作って復号 → The と the を最も多く含むものを探す」で暗号化キーを特定できました。

```
1  In[]:= Clear["Global`*"];
2  org = First@Import["0059_cipher.txt", "CSV"];
3  RepeatedTiming[
4   (* キーの候補を全部作る *)
5   keys = Tuples[
6     Range[First@ToCharacterCode@"a", First@ToCharacterCode@"z"], 3];
7  
8   (* 正しいキーを探す *)
9   countThe[key_] :=
10   SequenceCount[Flatten[BitXor[key, #] & /@ Partition[org, 3]],
11    Alternatives @@ (ToCharacterCode /@ {"The", "the"})];
12  encryptionKey = First@MaximalBy[keys, countThe@# &];
13 
14  (* 復号して解答 *)
15  decrypt[key_] :=
16   FromCharacterCode /@
17    Flatten[BitXor[key, #] & /@ Partition[org, 3]];
18  ans = First@Total[ToCharacterCode@decrypt@encryptionKey];
19  {ans, StringJoin@FromCharacterCode@encryptionKey}]
20 
21 Out[]= {1.37121, {129448, "exp"}}
```

`0059_cipher.txt` の中身は $1455 = 3 \times 485$ 個の数字のリストでした。10 行目と 17 行目ではこれを 3 個ごとに 3 つごとに区切って，キーとの XOR を取っています。

この部分は次のようにキーを 485 個並べたリストを作って XOR を取る方が速いのですが，後学のためにこうしました。

```
decrypt[key_] :=
  Module[{lst = Flatten@Table[key, Quotient[Length@org, 3]]},
   FromCharacterCode@BitXor[org, lst]];
countThe[key_] := StringCount[decrypt@key, {"the", "The"}];
```

暗号化鍵は「exp」でした。復号するとオイラーについての文章になります。

> An extract taken from the introduction of one of Euler's most celebrated papers, "De summis serierum reciprocarum" [On the sums of series of reciprocals]: I have recently found, quite unexpectedly, an elegant expression for the entire sum of this series 1 + 1/4 + 1/9 + 1/16 + etc., which depends on the quadrature of the circle, so that if the true sum of this series is obtained, from it at once the quadrature of the circle follows. Namely, I have found that the sum of this series is a sixth part of the square of the perimeter of the circle whose diameter is 1; or by putting the sum of this series equal to s, it has the ratio sqrt(6) multiplied by s to 1 of the perimeter to the diameter. I will soon show that the sum of this series to be approximately 1.644934066842264364; and from multiplying this number by six, and then taking the square root, the number 3.141592653589793238 is indeed produced, which expresses the perimeter of a circle whose diameter is 1. Following again the same steps by which I had arrived at this sum, I have discovered that the sum of the series 1 + 1/16 + 1/81 + 1/256 + 1/625 + etc. also depends on the quadrature of the circle. Namely, the sum of this multiplied by 90 gives the biquadrate (fourth power) of the circumference of the perimeter of a circle whose diameter is 1. And by similar reasoning I have likewise been able to determine the sums of the subsequent series in which the exponents are even numbers.