自主課題研究
ホモフォニック符号に対する区間アルゴリズム

名列番号 037 名前 中村 健二

平成 19 年 2 月 2 日

1 目的

ある入力列を別の出力列として変換する。このような暗号システムには、一様乱数が必要である。偏った乱数から一様乱数を導き出すアルゴリズムとして区間アルゴリズムを用いて、ホモフォニック符号を理解することを目的とする。

2 研究の流れ

乱数生成問題を区間アルゴリズムによりアプローチし、この区間アルゴリズムを利用して、ホモフォニック符号を理解する。

2.1 区間アルゴリズム

偏りのないコイン投げ X を何回か繰り返すことで、確率変数 Y を生成するとする。X を情報源として、情報源アルファベットを $\mathcal{X} = \{a, b, c\}$、$X$ の確率分布を $P(X = a) = 1/2, P(X = b) = 1/4, P(X = c) = 1/4$ とする。φ は、$\mathcal{X}$ から $\mathcal{Y}^*$($\mathcal{Y}^*$ は $\mathcal{Y}$ 上の有限ブロック全体の集合) への写像を次のように定める：

$$\varphi \ : \begin{cases} a \longmapsto & 0 \ , \\ b \longmapsto & 10 \ , \\ c \longmapsto & 11 \ . \end{cases}$$

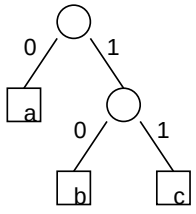

Fig1:符号木

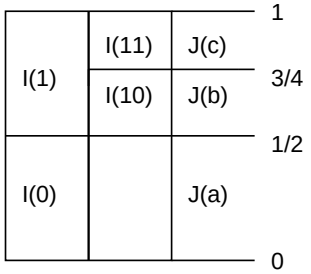

Fig2: 情報源の確率分布をp=(1/2,1/4,1/4), コードアルファベットをZ={0,1}, 配分比をq=(1/2,1/2) とする.

Fig1 に Intervsl Algorithm の**符号木**† を示す。

[**アルゴリズム**]

0) $s = \lambda$(空列) をセットする。$I(s) = [0, 1)$ とする;

1) もし、$I(s) \subseteq J(k)$ なら、k $(k \in \mathcal{X})$ を出力して終了する。

2) 確率分布$\boldsymbol{q}$に従い、$I(s)$ を M 個に分割する。ここで、$a \in \{1, 2, \cdots, M\}$ としたとき、$s = sa$ としてステップ 1 へいく。

2.2 ホモフォニック符号

X を情報源、$\mathcal{X} = \{a, b\}$ を情報源アルファベットとする。X の確率分布を $P(X = a) = 3/4, \ P(X = b) = 1/4$ とする。φ は、$\mathcal{X}$ から $\mathcal{Y}^*$($\mathcal{Y}^*$ は例 1 で定義した) へのランダム写像を次のように定める：

$$\varphi \ : \begin{cases} a \longmapsto \begin{cases} 0 & \text{with probability } 2/3, \\ 10 & \text{with probability } 1/3, \end{cases} \\ b \longmapsto \quad 11 \ \text{ with probability } \ 1. \end{cases}$$

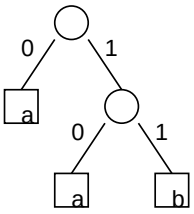

Fig3: 符号木

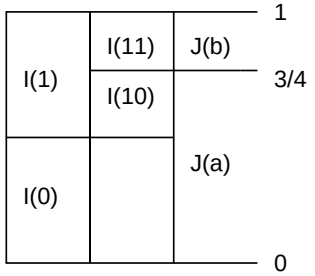

Fig4: 情報源の確率分布をp=(3/4,1/4), コードアルファベットをZ={0,1}, 配分比をq=(1/2,1/2) とする.

このランダム写像 φ は一つのホモフォニック符号を与える。Fig3 に φ の符号木を示す。このとき、文字 a は、符号語 0 に確率 $2/3$ で,10 に確率 $1/3$ でランダムに写像される。符号語 0 と 10 は情報源文字 a に対するホモフォーン（同音異義語）と呼ばれる。

[**アルゴリズム**]

0) 情報源の出力 $X = k \in \{1, 2, \cdots, N\}$ を読みとり、区間 $J(k)$ を計算し、$J(k)$ 上の均等分布乱数 r を選ぶ。

1) $s := \lambda$ (空列),
$I(s) := [0, 1)$ とする。

2) もし、$I(s) \subseteq J(k)$ ならば、k に対する符号語として記号列 s を出力する。そして、アルゴリズムを終らせる。

3) $I(s)$ を M 個の部分区間 $I(sj) \quad (j = 1, \cdots, M)$ に分割する。
$r \in I(sa)$ となるような a を見つける。
$s := sa$ とする。そしてステップ 2 へ戻る

3 まとめ

今回の研究によって、目的であるホモフォニック符号について、乱数を用いることで一意にターゲット乱数を生成できることがよく理解できた。