数理解析研究所講究録
第 421 巻 1981 年 218-231

自己シャフルに関する判定問題について

京産大 理　岩間一雄

1. まえがき

Σを有限アルファベットとする。自己シャフルSは，Σ^*からΣ^*の有限部分集合族の中への写像であり，

$$S(x) = x \odot x$$

で定義される。ここで，$\odot$はシャフル演算[1,2]であり，$x \in \Gamma^*$, $y \in \Delta^*$（ΓとΔは有限アルファベット）に対し，

$$x \odot y = \{x_1 y_1 \cdots x_n y_n \mid x_i \in \Gamma^*, y_i \in \Delta^*, x_1 \cdots x_n = x, y_1 \cdots y_n = y\}$$

で定義される。例えば，

$$ab \odot cd = \{abcd, acbd, acdb, cdab, cabd, cadb\}$$

である。S^{-1}を逆自己シャフルと呼び，

$$S^{-1}(y) = \{x \mid y \in S(x)\}$$

で定義する。

自己シャフルを考えるに至った動機は，通常のシャフルの場合[2]と同様，非同期並列処理系のモデル化に利用することを

めざしたものであった。[3] 一方，数学的には，自己シャッフルは記号列に対する一種の符号化とみなせる。この観点からみると，自己シャッフルは，その構造上の単純さにもかかわらず，相当奥深いものがあり，種々の組合せ論的問題を提供してくれる。例題を上げよう。

(1) 記号列 $x \in \Sigma^*$ が与えられたとき，$\{x\} \in S^{-1}(y)$ となる記号列 $y \in \Sigma^*$ を求める問題。（つまり，一意復号可能な y を $S(x)$ の中から選べという問題。）$x = 1010$ が与えられたとしよう。このとき，$y = 11001100$ は正答ではない。なぜなら，$S^{-1}(11001100) = \{1010, 1100\}$ であるから。$y = (10)^4$ は正答である。$x = 101010$ に対しては，$y = (10)^6$ は正答ではない。なぜなら，$S^{-1}((10)^6) = \{101010, 100110\}$ であるから。$y = 10110011001\,0$ は正答である。（この問題に関しては，$S(x)$ の中に一意復号可能な列が存在しないような $\{0, 1\}$ 上の列 x が存在することが判っている。[4]）

(2) $S^{-1}(000100000001000) = \phi$（空集合）であることに注意されたい。与えられた列 y に対し，$S^{-1}(y) = \phi$ かどうか判定する効率のよいアルゴリズムが存在するであろうか。

本稿では(2)の問題を扱い，$S^{-1}(y) = \phi$ かどうかを一般の y に対して判定する効率のよいアルゴリズムは存在しない，つまりアルゴリズム論の立場からは，"$S^{-1}(y) = \phi$?" は手に負

えない（intractable）問題[5]であることを示す。本結果の延長として，$S^{-1}(y)$に属す列を具体的に求める問題（つまり，自己シャフルによって符号化された列を復号する問題）も同様に手に負えない問題であることが示せる。このことは，自己シャフルが，暗号系に利用できる可能性を有していることを示唆している。

2. 逆自己シャフル空問題

アルファベットΣ上の列yに対し，$S^{-1}(y)$が空集合になるか否かを判定せよという問題を逆自己シャフル空問題と呼ぶ。本稿では，Σの大きさを制限しない（即ち，個々の例題として与えられる列yには最大$|y|$個の異った記号が出現してよい）一般の場合を扱う。Σの大きさを（例えば2値に）制限した場合については現在のところ未解決である。以下に二，三の例題を与える。

(1) $S^{-1}(aab)=\phi$ （∵ aab の長さが奇数である。）

(2) $S^{-1}(aaab)=\phi$ （∵ aが奇数回出現している。）

(3) $S^{-1}(abbbbbaabbba)=\phi$ （∵ $x\in S^{-1}(abbbbbaabbba)$ ならxはプレフィクスとして$abbbbb$ を有す。従って，xに現れるbの個数は少くとも5個であるから，$x\in S^{-1}(y)$ ならyにはbが少くとも10個現れる必要がある。）

このように，$S^{-1}(y)=\phi$ であるための簡単な十分条件はいくつか得られる。一般の場合に関しては，例えば以下の様な総当たり的手法が考えられる。与えられた列 y の長さを l としたとき，長さ $l/2$ の（記号が連続する必要のない）すべての部分列に対し，その部分列と，y からその部分列を除いた残りの部分列が一致するかどうか調べる。この方法によれば調べるべき場合の数は ${}_lC_{l/2}$ となり，l に関して指数オーダになる。しかし，この手法より本質的に効率のよいアルゴリズムは存在しない（と考えてよい[5]）というのが本稿での結果である。

〔定理〕　逆自己シャフル空問題はNP－完全である。

3. 証明

3.1　基本方針

一般に，問題QがNP－完全であることを証明するには，(i) QがクラスNPに属する（非決定性チューリング機械によって多項式時間で解が得られる）ことと，(ii)既にNP－完全であると判っている別の問題Q′がQへ多項式時間変換可能(polynomially transformable)であることを示せばよい。逆自己シャフル空問題に関しては，(i)は容易であるから(ii)のみ示す。問題Q′としては，3充足可能性問題(3SAT[5])を利用す

る。3SATは，3乗法標準形の論理式

$$A=(x'_{11}+x'_{12}+x'_{13})(x'_{21}+x'_{22}+x'_{23})\cdots\cdots(x'_{n1}+x'_{n2}+x'_{n3})$$

が充足可能であるかどうかを判定せよという問題である。ここで，x'_{ij} はリテラルと呼ばれ，論理変数 x_k 又はその否定 $\overline{x}_k$ である。ここでは，論理式Aの各和項の3個の変数はすべて異っているとの制限を設ける。つまり，$(x_k+x_k+x_h)$ とか，$(x_k+\overline{x}_k+x_h)$ といった項の現れる式は問題の対象にしない。この制限が3SATのNP-完全性に影響を与えないことは明らかである。

証明の目標は，3乗法標準形の上記条件を満たす任意の論理式Aを，条件

(＊) Aが充足可能であるための必要十分条件は $S^{-1}(y)\neq\phi$ であることである

を満たす記号列 y へ変換する多項式時間アルゴリズムが存在することを示すことである。以下，3.2では式Aを列 y へ変換する規則を与え，そのような変換が決定性多項式時間で実行可能であることを示す。3.3では，得られた列 y が上記条件(＊)を満足することを示す。

3.2 変換規則

論理式Aが n 項より成るとする。Aより変換して得られる

記号列yはおおよそ以下の形である。

$a_{00}\, I_{00}\, a_{00} a_{01}\, I_{01}\, a_{01}\cdot$

$a_{10}\, I_{10}\, a_{10}\;\; a_{11}\, I_{11}\, a_{11}\; \cdots\cdots\; a_{1q}\, I_{1q}\, a_{1q}\cdot$

$\vdots$

$a_{n0}\, I_{n0}\, a_{n0}\, a_{n1}\, I_{n1}\, a_{n1}\; \cdots\cdots\; a_{nq}\, I_{nq}\, a_{nq}\cdot$

$a_{n+1\,0}\, I_{n+1\,0}\;\; a_{n+1\,0}\;\; a_{n+1\,1}\;\; I_{n+1\,1}\; a_{n+1\,1}$

ここで，I_{ij} は以下で説明される記号列であり，$a_{00}, a_{01}, a_{10}, a_{11}, \cdots, a_{n+1\,0}, a_{n+1\,1}$ は，I_{ij} に現れずかつ互いに相異った記号である。I_{ij} に現れる記号は以下のアルファベットの中から選ばれる。

$\{1, 2, \cdots\cdots, m\} \cup$

$\{\#_1, \#_2, \cdots\cdots, \#_m\} \cup$

$\{\#'_1, \#'_2, \cdots, \#'_m\} \cup$

$\{*_1, *_2, \cdots\cdots, *_m\} \cup$

$\{\$_1, \$_2, \cdots\cdots, \$_{m+2}\}$

ここで，mは式Aに現れる異った論理変数の数である。

以下で述べる様に，$I_{00}, I_{01}, I_{n+1\,0}, I_{n+1\,1}$ は式Aに現れる変数の数mによって決まる。又，$I_{i0}, I_{i1}, \cdots, I_{iq}$ $(1 \leq i \leq n)$ に関しては，$I_{i0}, I_{i1}, I_{i2}, I_{i5}, I_{i6}, I_{i7}$については，変数の数$m$と式Aの左から$i$番目の項に現れるリテラルによって決まり，$I_{i3}, I_{i4}, I_{i8}, I_{iq}$ はmによってのみ決まる。説明を簡

単にするため，Aは5変数 x_1, x_2, x_3, x_4, x_5 を用い，又，A の最初の項は $x_2+\overline{x}_3+x_5$ であると仮定して，$I_{00}, I_{01}, I_{10}, \cdots, I_{19}, I_{n+1\,0}, I_{n+1\,1}$ がどのような列になるかを以下に与える。$I_{i0}, \cdots, I_{i9}$ $(i \geq 2)$ についても同様である。

$I_{00} = 12345$

$I_{01} = \#_1 1 *_1 1 \#'_1 \#_2 2 *_2 2 \#'_2 \#_3 3 *_3 3 \#'_3 \#_4 4 *_4 4 \#'_4 \#_5 5 *_5 5 \#'_5$

$I_{10} = H(x_1 \overline{x}_1 x_2 x_2 \overline{x}_2 x_3 \overline{x}_3 \overline{x}_3 x_4 \overline{x}_4 x_5 x_5 \overline{x}_5)$ （Hは，$H(x_j) = \#_j j *_j \#'_j$，$H(\overline{x}_j) = \#_j *_j j \#'_j$ で定まる準同型。$x_1 \overline{x}_1 x_2 x_2 \overline{x}_2 \cdots$ は，$x_1 \overline{x}_1 x_2 \overline{x}_2 x_3 \overline{x}_3 x_4 \overline{x}_4 x_5 \overline{x}_5$ に，Aのオ一項に現れるリテラル x_k を x_k と $\overline{x}_k$ の間にそう入して得られる。）

$I_{11} = \delta_1 \$_1 \delta_2 \$_2 \delta_2 \$_3 \delta_3 \$_4 \delta_3 \$_5 \delta_4 \$_6 \delta_5 \$_7 \delta_5$ （$\delta_j = \#_j *_j j j *_j \#'_j$ である。Aのオ一項にリテラル x_2，$\overline{x}_3$，x_5 が現れていることに対応して $\delta_2, \delta_3, \delta_5$ が二度づつ現れている。）

$I_{12} = I_{11}$

$I_{13} = H(x_1 \overline{x}_1 x_2 \overline{x}_2 x_3 \overline{x}_3 x_4 \overline{x}_4 x_5 \overline{x}_5) \$_1 H(x_1 \overline{x}_1 x_2 \overline{x}_2 x_3 \overline{x}_3 x_4 \overline{x}_4 x_5 \overline{x}_5) \$_2 H(x_1 \overline{x}_1 x_2 \overline{x}_2 x_3 \overline{x}_3 x_4 \overline{x}_4 x_5 \overline{x}_5)$

$I_{14} = I_{13}$，$I_{15} = I_{16} = I_{11}$，$I_{17} = I_{10}$

$I_{18} = I_{19} = \delta_1 \$_1 \delta_2 \$_2 \delta_3 \$_3 \delta_4 \$_4 \delta_5 \$_5$

$I_{n+1\,0} = I_{01}$，$I_{n+1\,1} = I_{00}$

一般化された変換規則を得ることは容易である。Aの項数をnとしたとき，yの長さはn^2のオーダになることに注意されたい。I_{ij}の構成規則はいずれも簡単なものであり，Aからyへの変換が多項式時間で実行できることは説明を要しないであろう。

3.3　変換規則の正統性

前節の規則によって式Aより得られた記号列yが3.1で述べた条件(*)を満たしていることを証明する訳であるが，本稿では，その基本的考え方を中心にして簡潔に述べておくにとどめる。

先ず，式Aが充足可能であると仮定して，このとき，3.2で得られた列yが等しい2つの列z_1とz_2に分解可能であること（以下，$y \in z_1 \odot z_2$であることを，yはz_1とz_2に分解可能であると呼ぶことにする）を示そう。なお，3.2の仮定（式Aは5変数を用い，最初の項は$x_2+\bar{x}_3+x_5$である）を継続して説明を進める。

(1)　列yがz_1とz_2（$z_1=z_2$）に分解可能であるなら，列yに記号a_{ij}がちょうど2回づつ現れていることに注目すると，z_1とz_2は必ず以下の形になることが容易に判る（図1参照）。

$$z_1 = a_{00}\alpha_{00}\,a_{01}\alpha_{01}\,a_{10}\alpha_{10}\,a_{11}\cdots\cdots a_{n+1\,0}\alpha_{n+1\,0}\,a_{n+1\,1}$$

$$z_2 = a_{00}\beta_{00}\,a_{01}\beta_{01}\,a_{10}\beta_{10}\,a_{11}\cdots\cdots a_{n+1\,0}\beta_{n+1\,0}\,a_{n+1\,1}$$

ここで,

(i) $\alpha_{00}=\beta_{00}$, $\alpha_{01}=\beta_{01}$, $\alpha_{10}=\beta_{10}$, ……, $\alpha_{n+1\,0}=\beta_{n+1\,0}$

(ii) $\alpha_{00}=I_{00}$, $I_{01}\in\beta_{00}\odot\alpha_{01}$, $I_{10}\in\beta_{01}\odot\alpha_{10}$, $I_{11}\in\beta_{10}\odot\alpha_{11}$, ……

…… $I_{n+1\,0}\in\beta_{nq}\odot\alpha_{n+1\,0}$, $\beta_{n+1\,0}=I_{n+1\,1}$

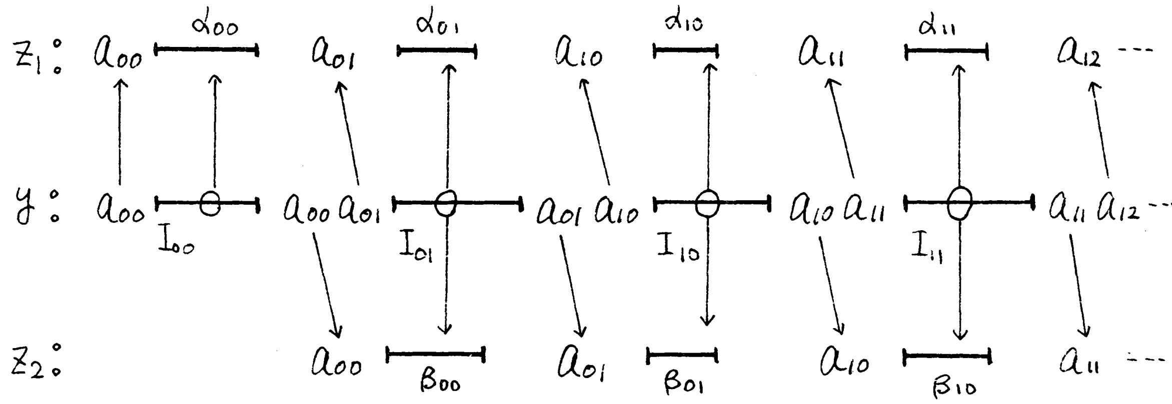

図1　列 y の z_1 と z_2 への分解

(2) 列 y を z_1 と z_2 に分解しようとするとき, (1)より

$$\alpha_{00}=\beta_{00}=I_{00}=12345$$

は自動的に定まる。そこで次は,

$$I_{01}=\#_1 1 *_1 1 \#'_1 \#_2 2 *_2 2 \#'_2 \#_3 3 *_3 3 \#'_3 \#_4 4 *_4 4 \#'_4 \#_5 5 *_5 5 \#'_5$$

を, $\beta_{00}=12345$ と α_{01} に分解することになる。α_{01} としては, 明らかに 2^5 通りの異った列が可能である。Aは充足可能であるとの仮定より, Aのすべての項の少くとも1個のリテ

ラルの値を1にするような変数への真理値0，1の割当てが存在する。仮に，

$$x_1 = x_4 = x_5 = 0 \ , \quad x_2 = x_3 = 1$$

がそのような割当てであるとする。このとき我々はα_{01}として，

$$\alpha_{01} = \#_1 *_1 1 \#_1' \#_2 2 *_2 \#_2' \#_3 3 *_3 \#_3' \#_4 *_4 4 \#_4' \#_5 *_5 5 \#_5'$$
$$= H(\overline{x}_1 x_2 x_3 \overline{x}_4 \overline{x}_5)$$

を選択する。

(3) 次に，

$$I_{10} = H(x_1 \overline{x}_1 x_2 x_2 \overline{x}_2 x_3 \overline{x}_3 \overline{x}_3 x_4 \overline{x}_4 x_5 x_5 \overline{x}_5)$$

を，$\beta_{01} = \alpha_{01} = H(\overline{x}_1 x_2 x_3 \overline{x}_4 \overline{x}_5)$ と α_{10} に分解する。α_{10}として，

$$\alpha_{10} = H(x_1 x_2 \overline{x}_2 \overline{x}_3 \overline{x}_3 x_4 x_5 x_5)$$

が選択できることは明らかである。ここで，$x_1 = x_4 = x_5 = 0$，$x_2 = x_3 = 1$ という割当てが，式Aの才1項の3個のリテラルのうち，$\overline{x}_3$とx_5を0にしていることに対応し，α_{10}には，$H(\overline{x}_3\overline{x}_3)$ と $H(x_5 x_5)$という，同じリテラルが2個連続したものをHで写した部分列が出現していることに注意しよう。なお，ここでは，α_{10}が準同型Hの値域に納まる（つまり，$H(\alpha_{10}') = \alpha_{10}$ となる列α_{10}'が存在する）ようにI_{10}を分解している。以下，このような分解を標準分解と呼ぶことにする。もちろん，標準分解でない分解も可能である。

(4) I_{11}とI_{12}の役割は後に述べる。I_{11}を$\beta_{10} = \alpha_{10}$とα_{11}，

I_{12} を $\beta_{11}=\alpha_{11}$ と α_{12} に分解していくとき，α_{10} が I_{10} の標準分解によって得られたものであるなら，

$$\alpha_{12}=\alpha_{10}$$

となる分解しか存在しないことが容易に判る。

(5) I_{13} を $\beta_{12}=\alpha_{12}=H(x_1x_2\overline{x}_2\overline{x}_3\overline{x}_3x_4x_5x_5)$ と α_{13} に分解する。ここでも標準分解を採用するので，準同型 H を無視して，

$$I'_{13}=\underline{x_1}\overline{x}_1\underline{x_2}\underline{\overline{x}_2}x_3\underline{\overline{x}_3}x_4\overline{x}_4x_5\overline{x}_5\$_1x_1\overline{x}_1x_2\overline{x}_2x_3\underline{\overline{x}_3}\underline{x_4}\overline{x}_4$$
$$\underline{x_5}\overline{x}_5\$_2x_1\overline{x}_1x_2\overline{x}_2x_3\overline{x}_3x_4\overline{x}_4\underline{x_5}\overline{x}_5$$

より，$\alpha'_{12}=x_1x_2\overline{x}_2\overline{x}_3\overline{x}_3x_4x_5x_5$ を抜き出すことを考えればよい。抜き出し方は，上記下線の様に一意に決まる。注意すべきことは，α'_{12} において，同じリテラルが連続するたびに，$\$_1$ あるいは $\$_2$ を飛び越して抜き出さねばならないことである。もし，α'_{12} に同じリテラルの連続が3ヶ所に現れる（すべてのリテラルの値を0にする）ならこのような抜き出し方はもはや不可能になる。（なお，このことは，I_{13} の分解が標準分解であるか否かにはよらない。）そうして，

$$\alpha_{13}=H(\overline{x}_1x_3x_4\overline{x}_4x_5\overline{x}_5)\$_1H(x_1\overline{x}_1x_2\overline{x}_2x_3\overline{x}_4\overline{x}_5)\$_2$$
$$H(x_1\overline{x}_1x_2\overline{x}_2x_3\overline{x}_3x_4\overline{x}_4\overline{x}_5)$$

が得られる。なお，仮に α'_{12} において同じリテラルの連続が1ヶ所以下であった場合，上記の様な抜き出し方は一通りで

はない。この場合，いずれでもよいから適当に一つの抜き出し方を選択すればよい。

(6) I_{14}～I_{19} の分解については詳細な説明は省く。I_{14} と I_{17} の分解を標準分解で行うなら分解法はほぼ一意であり，

$$\alpha_{19} = H(\bar{x}_1 x_2 x_3 \bar{x}_4 \bar{x}_5)$$

が得られる。こうして(2)で選択した α_{01} が再現された。I_{14}～I_{19} を設けた目的は α_{01} を再現することである。

以上整理して述べると，(i) α_{01} としては 2^5 通りの異った列が可能であり，それぞれの列は，$x_1=0, x_2=0, \cdots, x_5=0$ から $x_1=1, x_2=1, \cdots, x_5=1$ までの 2^5 通りの真理値割当のうちのいずれかに一対一で対応している。(ii) I_{10}～I_{19} を設けた目的は，(i)で選ばれた α_{01} に対応する真理値割当てが式 A の才1項を1にするかどうかを調べることである。そのためには，I_{10} と I_{13} の役割が特に重要であることが理解されるであろう。(iii) もしその真理値割当てが才1項を1にしているなら，I_{10}～I_{19} を上記の方針で分解することが可能であり，α_{19} として先に選択した α_{01} が再現される。(iv) この α_{19} を利用し，I_{20}～I_{29} によって，α_{01} に対応する真理値割当が式 A の才2項を1にしているかどうかを調べる。I_{30} 以降も同様である。

こうして，式 A が充足可能であるなら列 y が z_1 と z_2 (z_1 =

z_2）に分解できることが判った。逆に，式Aが充足可能でないなら列yが等しい２つの列に分解できないことを示す。

(7) Aは充足可能でないから，変数にどのような真理値割当てを行っても，Aの少くとも１つの項のすべてのリテラル値を０にする。仮に，ある列をα_{01}として選択し，その列に対応する真理値割当がAの第i項で初めてすべてのリテラル値を０にしたと仮定する。又，$j \leqq i-1$であるすべてのjに対し，I_{j0}, I_{j3}, I_{j4}, I_{j7}の分解に標準分解を適用したと仮定する。上記議論より，$\alpha_{i-1\,q}$として最初に選んだα_{01}が再現されていることが判る。さらにI_{i0}の分解も標準分解を適用するなら，上記(5)で述べたようにI_{i3}の分解が不可能になってしまう。

このように，もし式Aが充足可能なら列yの等しい列z_1とz_2への分解が可能であり（$S^{-1}(y) \neq \phi$），Aが充足可能でないなら列yのz_1とz_2への分解は，I_{k0}, I_{k3}, I_{k4}, I_{k7}の分解に標準分解を適用するという条件のもとでは不可能であることが判った。従って，変換規則の正統性の証明を完成するためには，列yがz_1とz_2に分解できるためには，式Aの充足性には関係なく，I_{k0}, I_{k3}, I_{k4}, I_{k7}の分解には標準分解を用いることが必要条件であるということをいえばよい。そのためには，もし標準分解を採用しないとしたら，I_{k1}, I_{k5}, I_{k8}

の分解が不可能になることを示せばよい。この証明は多くの場合分けを必要とするがそれ程困難ではない（省略）。

文献

(1) S. Ginsburg and E.H. Spanier, "Mapping of Languages by Two-Tape Devices," JACM, 12, 423-434, 1965.

(2) A.C. Shaw, "Software Descriptions with Flow Expressions," IEEE Trans. Software Eng., SE-4, 242-254, 1978.

(3) 岩間, 上林, "自己シャッフルされた記号列を入力とする有限オートマトンについて," 京大数解研講究録, 381, 207-222, 昭55年4月.

(4) 岩間, "記号列の自己シャッフルについて," 信学技報, AL80-22, 昭55年7月.

(5) A.V. Aho, J.E. Hopcroft and J.D. Ullman, The Design and Analysis of Computer Algorithms, Addison-Wesley, 1974.