数理解析研究所講究録
第454巻1982年45-60

無限列の圧縮可能性

東工大　理学部　　小林孝次郎

1．　はじめに　αを，i番目のビット$\alpha(i)$が次の規則で定義されるような2進無限列（以後単に無限列と呼ぶ）とする．

$$\alpha(i) = \begin{cases} 1 & i\text{番目のTuring機械（あるいはFORTRANプログラム）にある決まった入力（例えば0）を与えると停止するとき，} \\ 0 & \text{そうでないとき．} \end{cases}$$

よく知られているように，αは計算可能ではない．その意味で，αは我々人間にとって，神の啓示によってのみ与えられる一種の聖なるメッセージであるといえる．しかしαに含まれるすべてのビットが，神の啓示を必要とする不可知の情報であるわけではない．

例えば，n_0，n_1，n_2，…を次のようなFORTRANプログラムの番号としよう．

```
I = 0
STOP
END
```

```
I = 0
I = 0
STOP
END
```

```
I = 0
I = 0
I = 0
STOP
END
```

…

これらのプログラムすべてが停止することは容易にわかる．従って $\alpha(n_0)$, $\alpha(n_1)$, $\alpha(n_2)$, … の値はすべて1である．つまりαのビットのうちのいくつかは，我々にも知り得るのである．

もう1つの例をあげよう．今任意のFORTRAN プログラムが与えられたものとし，その番号をnとする．そのプログラムはSTOP文を含んでいると仮定する．そのSTOP文を2つのSTOP文，3つのSTOP文，… 等で置きかえて得られる次のようなプログラムを考え，それの番号をm_0, m_1, … とする．

```
  ⋮
STOP
  ⋮
```

（もとのプログラム）

```
  ⋮
STOP
STOP
  ⋮
```

```
  ⋮
STOP
STOP
STOP
  ⋮
```

…

我々は$\alpha(n)$の値は知らないかも知れないが，$\alpha(n)$，$\alpha(m_0)$，$\alpha(m_1)$，… の値すべてが等しいことは容易にわかる．従って，無限個のビット$\alpha(n)$，$\alpha(m_0)$，$\alpha(m_1)$，… のうちの1ビットの値がわかるだけで十分である．このことから特に，αを印刷した聖典から有限ビットが虫くいによって失なわれても，われわれはそれを修復することができることがわかる．

以上のような意味で，無限列αは冗長性を持っている．ところで有限列が冗長性を持つ場合にはそれを利用してより短い有限列で表現することができた．我々は，無限列の場合にもそのようなことが可能であるか，という問題を考える．

2．基本概念　fを自然数から自然数への関数とする．無限列αは，次の条件を満足するようなアルゴリズム$\mathscr{A}$と無限列βが存在するときf圧縮可能であるという：任意のnに対し，$\mathscr{A}$はβの最初の$f(n)$ビット$\beta(1)$，$\beta(2)$，…，$\beta(f(n))$より$\alpha(n)$の値を計算する．もう少し正確な定義は次の通りである．

M_0，M_1，M_2，… を，無限列をoracleとして使うTuring機械の，ある自然な数え上げとする．任意のiと無限列βに対し，NからNへの部分関数$\psi_{i,f}{}^{\beta}$を次のように定義する．M_iにoracleとしてβ，入力としてnを与えたとき，M_iが

βの最初の$f(n)$ビットの部分以外の部分を見ないである値mを出力して停止する場合には$\psi_{i,f}{}^{\beta}(n)=m$である．それ以外の場合には$\psi_{i,f}{}^{\beta}(n)$の値は定義されない．$\alpha$は，$\alpha=\psi_{i,f}^{\beta}$なる$i$，$\beta$が存在するとき，$f$<u>圧縮可能</u>（$f$ compressible）であるという．

圧縮可能性とKolmogorov，Chaitin，Martin-Löf，Daley，Schnorr，Lovelandらによって研究されたrandomnessの間には，強い関係（類似性といくつかの本質的な差）がある．このことについては［1］を参照されたい．

3．いくつかの基本的な結果　本節では圧縮可能性に関するいくつかの基本的な結果を要約する．

まずαが$f(n)$圧縮可能なら，任意の常数kに対しαは$f(n)-k$圧縮可能である．（αの最初のkビットを，アルゴリズムの中に繰りこんで記憶しておけばよい．）これは‘constant compressibility theorem’とでもいうべきもので，有用な道具である．

このことから直ちに，$n-f(n)$が上に有界なら（つまり任意のnに対し$n-f(n)\leq k$なる常数kが存在すれば），任意のαはf圧縮可能であることがでてくる．従って，我々に興味があるのは，$n-f(n)$が上に有界でない場合のみで

ある．

$n-f(n)$ が上に有界でない場合には，‘ほとんどすべて’の α が f 圧縮可能でないことを示すことができる．

［定理1］　$n-f(n)$ が上に有界でない場合には，集合 $\{\alpha \mid \alpha$ は f 圧縮可能$\}$ は測度0のある集合の部分集合になっている．

ここで，すべての無限列の集合 Σ^∞ は2点集合 $\Sigma=\{0,1\}$ の無限積 $\Sigma\times\Sigma\times\cdots$ であると考え，Σ^∞ には Σ の測度 μ'（$\mu'(\{0\})=\mu'(\{1\})=1/2$）の積測度 μ が入れられているものと考えている．上の定理の証明は容易であるので省略する．

次に我々は，無限列のいろいろなクラス $\mathcal{C}$ といろいろな関数 f に対し

$$\alpha\in\mathcal{C} \Longrightarrow \alpha \text{ は } f \text{ 圧縮可能} \qquad (*)$$

というかたちの定理がなりたつかどうかを考えてみる．特に $\mathcal{C}$ が，集合 $\{i \mid \alpha(i)=1\}$ の算術的階層（arithmetical hierarchy）によって特徴づけられる場合に興味がある．

集合 $\{i \mid \alpha(i)=1\}$ が帰納的集合（recursive set）であるとき，α は<u>帰納的</u>であるという．その集合が帰納的に枚挙可能な集合（recursively enumerable set）であるとき，α は<u>帰納的に枚挙可能</u>である（r. e. と略する）という．

同じ集合が算術的階層のΔ_2に入っているとき，つまり

$$i \in \{ i \mid \alpha(i) = 1 \}$$

$$\leftrightarrow \exists x \forall y \; R(i, x, y) \leftrightarrow \forall x \exists y \; R'(i, x, y)$$

がなりたつような帰納的述語（recursive predicate）R，R'が存在するとき，αは<u>Δ_2無限列</u>であるという．明らかに

$$\alpha \text{が帰納的} \Rightarrow \alpha \text{が r.e.} \Rightarrow \alpha \text{が} \Delta_2 \text{無限列}$$

がなりたつ。つまり，帰納的なαの集合，r.e.なαの集合，Δ_2無限列の集合をそれぞれ$\mathcal{C}_{rec}$，$\mathcal{C}_{r.e.}$，$\mathcal{C}_{\Delta_2}$で表わすと，

$$\mathcal{C}_{rec} \subseteq \mathcal{C}_{r.e.} \subseteq \mathcal{C}_{\Delta_2}$$

である．

αが帰納的ならαは0-圧縮可能である．（αは，oracleを全く見ないで計算できる．）従って，$\mathcal{C} = \mathcal{C}_{rec}$なら，どんな$f$に対しても上記の式（*）はなりたつ．一方$\mathcal{C}_{\Delta_2}$に関しては，次の定理がなりたつ．

［定理２］　fが帰納的関数で$n - f(n)$が上に有界でない場合には，f圧縮可能でない無限列が$\mathcal{C}_{\Delta_2}$内に存在する．

この定理は，単純な対角線論法で証明することができる．この定理は，fが帰納的関数の場合には，全く自明な場合（$n - f(n)$が上に有界である場合）以外には上記の式（*）はなりたたないことを示している．

そこで残されている問題は，$\mathcal{C} = \mathcal{C}_{r.e.}$の場合に，どんな

fに対し上記の式（＊）が定理になりうるか，ということである．次節でこの問題を考える．

4．　r.e.な無限列の圧縮可能性　我々は本節で，fに関するある付帯条件のもとで，$\sum_{i=0}^{\infty} 2^{-f(i)}$が収束するとき，そのときに限り

$$\alpha \text{が r.e.} \Rightarrow \alpha \text{は} f \text{圧縮可能}$$

がなりたつことを示す．このことから，例えば$f(n) = n/2$，$\sqrt{n}$，$2 \operatorname{Log} n$，$\operatorname{Log} n + 2 \operatorname{Log} \operatorname{Log} n$などについては上のことがなりたち，$f(n) = \operatorname{Log} n$，$\operatorname{Log} n + \operatorname{Log} \operatorname{Log} n$などについては上のことがなりたたないことがわかる．（本稿では，特に指定しない限りLogの底は2である．）

［定理3］　fが$\sum_{i=0}^{\infty} 2^{-f(i)} < \infty$なる帰納的関数なら，任意のr.e.な無限列$\alpha$は$f$圧縮可能である．

（証明）　αをr.e.な無限列とする．$\sum_{i=0}^{\infty} 2^{-f(i)}$は収束するから$\sum_{\substack{1 \leq i \\ \alpha(i)=1}} 2^{-f(i)}$なる値（つまり，$\alpha(i) = 1$なる$i$（$1 \leq i$）に対する$2^{-f(i)}$の和）は有限のある値である．それを$\Omega$とする．$\Omega$の2進表現から小数点を除いて得られる無限列を$\beta$とする．ただし$\Omega < 1$なら$\beta$は小数点の右の位置からはじまるものとし，$\Omega \geq 1$なら$\beta$は1ではじまるように選ぶものとする．$\beta$のビットのうち，小数点より左のものの数を$n_0$

とする．

つまり，もしΩの2進表現が

$$0.0010110111 0110\cdots$$

ならβは

$$0010110111 0110\cdots$$

で $n_0 = 0$ であり，もしΩの2進表現が

$$1101.0011101001 10\cdots$$

ならβは

$$11010011101001 10\cdots$$

で $n_0 = 4$ である．あるところから後0が続く表現とあるところから後1が続く表現のどちらもが可能な場合（つまりΩが整数を2のベキ乗で割ったかたちの場合）は，どちらを採用してもよい．

我々は以下で，βの最初の $n_0 + f(n) + 2$ ビットから $\alpha(n)$ を計算できることを示す．このことからαが $n_0 + f(n) + 2$ 圧縮可能，従って constant compressibility theorem により $f(n)$ 圧縮可能であることがでてくる．

βの最初の $n_0 + f(n) + 2$ ビットの後に all 0を追加した無限列が2進表現であるような値を Ω_1，all 1を追加した無限列が2進表現であるような値を Ω_2 とする．（小数点はもとの位置を採用する．）Ω_1，Ω_2 は有理数で，その値はβの最

初の $n_0 + f(n) + 2$ ビットから計算できる．さらに

$$\Omega_1 \leq \Omega \leq \Omega_2,$$

$$\Omega_2 - \Omega_1 = 2^{-f(n)-2} = 2^{-f(n)}/4$$

がなりたつ．

次に，Ω_3，Ω_4 を次の式で定義する．

$$\Omega_3 = \Omega_1 - 2^{-f(n)}/8,$$

$$\Omega_4 = \Omega_2 + 2^{-f(n)}/8.$$

そうすると，この Ω_3，Ω_4 に対し

$$\Omega_3 < \Omega < \Omega_4,$$

$$\Omega_4 - \Omega_3 = 2^{-f(n)}/2$$

がなりたつ．

$\alpha(n)$ の値を決めるには，次のようにすればよい．我々は集合 $\{i \mid \alpha(i) = 1\}$ の要素を次々と enumerate してゆく．α は r.e. であるから，これは可能である．同時に，どの時点 t においても，その時点までに enumerate した $\{i \mid \alpha(i) = 1\}$ の要素に対する $2^{-f(i)}$ の和 $\Omega^{(t)}$ を計算しておく．この $\Omega^{(t)}$ に対しては

$$\Omega^{(t)} \leq \Omega,$$

$$\lim_{t \to \infty} \Omega^{(t)} = \Omega$$

がなりたつ．

もし $\alpha(n) = 1$ なら，n はいつか enumeration の中にで

てくるから，必ず我々はそのことを知る．

$\alpha(n)=0$ なら n は決して enumeration にでてこない．しかも $\Omega_3<\Omega=\lim_{t\to\infty}\Omega^{(t)}$ であるから，n が enumeration にでてこないまま $\Omega_3<\Omega^{(t)}$ となる時刻 t がある．このとき我々は $\alpha(n)=0$ であることを知る．なぜなら，$\alpha(n)=1$ なら次のような矛盾が得られるからである．

$$\Omega^{(t)}+2^{-f(n)}\leq\Omega<\Omega_4=\Omega_3+2^{-f(n)}/2$$

$$\therefore\quad \Omega^{(t)}<\Omega_3-2^{-f(n)}/2<\Omega_3 .$$

（証明終り）

Ω は G. Chaitin によって導入された（［２］）．Ω の esoteric な解釈については文献［３］を見られたい．文献［４］は上記文献［３］の紹介を含んでいる．

次の定理に進む前に，若干の記号を導入し補助定理を証明する．

f を，上に有界でない単調非減少関数とする．このとき $f^{-1}(m)$ を

$$f^{-1}(m)=\max\{n\mid f(n)\leq m\}$$

によって定義する．$f^{-1}(m)$ はある値以上のすべての m に対し定義され，f^{-1} 自身上に有界でない非減少関数である．さらに f が帰納的関数なら f^{-1} も帰納的部分関数である．

〔補助定理 1〕 fを上に有界でない単調非減少関数とし、m_0を$f^{-1}(m)$がすべての$m \geq m_0$に対し定義されている$\overbrace{\text{ような値}}^{m\text{の最小}}$とする。このとき $\sum_{i=0}^{\infty} 2^{-f(i)}$ が発散すれば $\sum_{m=m_0+1}^{\infty} 2^{-m}(f^{-1}(m) - f^{-1}(m-1))$ も発散する。

（証明） 集合Aの要素の数を#Aで表わすことにすれば、

$$
\begin{aligned}
&\sum_{i=0}^{\infty} 2^{-f(i)} \\
&= \sum_{m=m_0}^{\infty} 2^{-m} \cdot \#\{i \mid f(i) = m\} \\
&= \sum_{m=m_0+1}^{\infty} 2^{-m} \cdot \#\{i \mid f(i) = m\} \\
&\qquad\qquad + 2^{-m_0} \cdot \#\{i \mid f(i) = m_0\} \\
&= \sum_{m=m_0+1}^{\infty} 2^{-m} (f^{-1}(m) - f^{-1}(m-1)) \\
&\qquad\qquad + 2^{-m_0} \cdot \#\{i \mid f(i) = m_0\}
\end{aligned}
$$

がなりたつ。このことより明らか。

（証明終り）

【定理 4】 fを、上に有界でない単調非減少な帰納的関数とする。このとき、$\sum_{i=0}^{\infty} 2^{-f(i)}$ が発散すれば、f圧縮可能でないr.e.な無限列αが存在する。

（証明） 目的は

$$\forall i\ \forall \beta\ [\alpha \neq \psi_{i,f}^{\ \beta}]$$

がなりたつような、r.e.なαを作ることである。そのために、i, βに関する対角線論法を使う。今、αの値をあるところ

まで定義して

$$\forall \beta\ [\alpha \neq \psi_{0,f}{}^{\beta}],$$

$$\forall \beta\ [\alpha \neq \psi_{1,f}{}^{\beta}],$$

$$\vdots$$

$$\forall \beta\ [\alpha \neq \psi_{i-1,f}{}^{\beta}]$$

がなりたつようにできたとする．私々は，

$$\forall \beta\ [\alpha \neq \psi_{i,f}{}^{\beta}]$$

がなりたつようにするには，α をどう定義すればよいかを考える．

各 m に対し，集合 I_m を $I_m = \{ n \mid f^{-1}(m-1) < n \leq f^{-1}(m) \}$ によって定義する．$\alpha(n)$ の値は，$I_j \cup I_{j+1} \cup I_{j+2} \cup \cdots$ に含まれる n に対しては定義されていないものとする．

n を I_j に含まれるある値，u を長さ j の 0，1 のある有限列とし，$\alpha(n)$ の値を次のように定義してみる．

$$\alpha(n) = \begin{cases} 1 & \psi_{i,f}{}^{u000\cdots}(n) = 0 \text{ のとき}, \\ 0 & \text{そうでないとき}. \end{cases}$$

そうすると，u の拡張になっているような β（つまり，$\beta = u\beta'$ と表わせる β）に対しては，

$$\begin{aligned} \alpha(n) &\neq \psi_{i,f}{}^{u000\cdots}(n) \\ &= \psi_{i,f}{}^{\beta}(n) \end{aligned}$$

であるから，$\alpha \neq \psi_{i,f}{}^{\beta}$ がなりたつ．（ここで，$n \in I_j$ で

あるから $f(n) = j =$ ‘u の長さ’であることを使った．）

ところで，u の拡張となっているβの集合の測度は 2^{-j} である．従って，I_j に含まれる1個の n に対しうまく $\alpha(n)$ の値を定義することにより，測度が 2^{-j} のある集合に含まれるすべてのβに対し，$\alpha \neq \psi_{i,f}{}^{\beta}$ がなりたつようにできる．（しかもその集合は，長さ j のある 0，1 の有限列のすべての拡張の集合，という極めて簡単なかたちをしている．）

このことから，I_j に含まれるすべての n（それは $f^{-1}(j) - f^{-1}(j-1)$ 個ある）に対しうまく $\alpha(n)$ の値を定義することにより，測度が $2^{-j}\cdot(f^{-1}(j) - f^{-1}(j-1))$ のある集合に含まれるすべてのβに対し，$\alpha \neq \psi_{i,f}{}^{\beta}$ がなりたつようにできる．このことを I_{j+1}，I_{j+2}，… についてもくり返してゆけば，$I_j \cup I_{j+1} \cup \cdots \cup I_{j'}$ に含まれるすべての n に対しうまく $\alpha(n)$ の値を定義することにより，測度が

$$\sum_{m=j}^{j'} 2^{-m}\cdot(f^{-1}(m) - f^{-1}(m-1))$$

のある集合に含まれるすべてのβに対し，$\alpha \neq \psi_{i,f}{}^{\beta}$ がなりたつようにできる．

ところで補助定理1により，上の値は j' を大きくするといくらでも大きくなる．このことは，ある j' の段階で，$\alpha \neq \psi_{i,f}{}^{\beta}$ であるようなβの集合 X の測度が1になることを意味する．しかも X は，‘u のすべての拡張の集合’というかたち

の集合（uは0，1のある有限列）の有限個の和（しかも，互に素なものの和）になっている．従って，Xはすべての無限列の集合である．

これで，αの値をうまく定義して

$$\forall \beta\ [\alpha \neq \psi_{i,f}^{\beta}]$$

がなりたつようにできることがわかった．αがr.e.であることは，αの作り方より明らかである．

（証明終り）

［系1］　fを，上に有界でない単調非減少な帰納的関数とする．このとき，次の2つの条件は同値である．

（1）$\sum_{i=0}^{\infty} 2^{-f(i)}$が収束する．

（2）すべてのr.e.なαがf圧縮可能である．

この系を，具体的なfに対して適用すると，次の結果を得る．

［系2］　すべての整数$k \geq 0$と実数$\varepsilon > 0$に対し，任意のr.e.なαは$\mathrm{Log}\, n + \mathrm{Log}^2 n + \cdots + \mathrm{Log}^k n + (1+\epsilon)\,\mathrm{Log}^{k+1} n$圧縮可能である．すべての整数$k \geq 0$に対し，$\mathrm{Log}\, n + \mathrm{Log}^2 n + \cdots + \mathrm{Log}^k n + \mathrm{Log}^{k+1} n$圧縮可能でない，r.e.な$\alpha$が存在する．

（証明）　$k \geq 0$，$\varepsilon \geq 0$に対し$f_{k,\varepsilon}$を次のように定義する．

$$f_{k,\varepsilon}(n) = \log n + \log^2 n + \cdots + \log^k n + (1+\varepsilon)\log^{k+1} n .$$

この系は，次の式と系1よりすぐでてくる．

$$\sum_{i=i_0}^{\infty} 2^{-f_{k,\varepsilon}(i)}$$

$$= \sum_{i=i_0}^{\infty} \frac{1}{n(\log n)(\log^2 n)\cdots(\log^{k-1} n)(\log^k n)^{1+\varepsilon}}$$

$$\fallingdotseq \int_{i_0}^{\infty} \frac{dx}{x(\log x)(\log^2 x)\cdots(\log^{k-1} x)(\log^k n)^{1+\varepsilon}}$$

$$= \frac{1}{(\log e)^k} \int_{\log^k i_0}^{\infty} \frac{dy}{y^{1+\varepsilon}}$$

$$\begin{cases} < \infty & \varepsilon > 0 \text{ のとき}, \\ = \infty & \varepsilon = 0 \text{ のとき}. \end{cases}$$

ただしここで i_0 は，$\log^{k+1} n$ が定義されるような n の最小値である．（証明終り）

文献

1. K. Kobayashi, On compressibility of infinite sequences, Tokyo Institute of Technology, Dept. of Information Sciences, Research Reports on Information Sciences, No. C-34, March 1981.
2. G. Chaitin, A theory of program size formally identical to information theory, JACM 22

(1975), 329 - 340.

3. C. Bennett, On random and hard-to-describe numbers, IBM Research Report, RC 7483, May 1979.

4. M.ガードナー，乱数オメガの値を知れば宇宙の神秘も解明される，サイエンス，1980年1月．（文献［3］の紹介を含んでいる．）