Title	アルゴリズム的コード化定理を満たすChaitnマシンの特徴付けの研究
Author(s)	天谷，良章
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1479
Rights	
Description	Supervisor:石原 哉，情報科学研究科，修士

アルゴリズム的コード化定理を満たす Chaitinマシンの特徴付けの研究

天谷 良章

北陸先端科学技術大学院大学 情報科学研究科

2001年2月15日

キーワード: Chaitin マシン,Prefix コード, プログラムサイズの複雑さ, アルゴリズム的コード化定理, 最小プログラム.

1 背景

アルゴリズム的情報理論 (AIT) は、Turing の計算理論と Shannon の情報理論を組み合わせてできた結果である。その基本的概念はある対象を計算する最小プログラムのビットサイズでその対象の複雑さを測ることである。この概念は Kolmogorov 記述量やプログラムサイズの複雑さなどと呼ばれる。この分野は 1960 年代の半ばに A.N.Kolmogorov, R.J.Solomonoff, G.J.Chaitin によって独立に導入された。

さらに、1970 年代に入り、L.A.Levin, G.J.Chaitin によって AIT_2 (Prefix Complexity) が導入された。その戦略は、prefix-free 集合を定義域として持つ (Chaitin) マシンによってプログラムサイズの複雑さを測ることにある。これは、きれいな数学的性質を持っており、そのために AIT においては、標準的になりつつある。

AIT において重要な結果の一つとしてアルゴリズム的コード化定理がある。アルゴリズム的コード化定理によると、任意の万能 Chaitin マシンの記述量 (複雑さ) は、そのアルゴリズム的確率のエントロピーに関して漸近的に最適 (ある定数で抑えられる) であることが知られている。

この結果をより一般的にするために、必ずしも万能ではないマシンのクラスや任意の semi-distributions、そして未知な定数に興味が持たれる。そこで、本研究ではアルゴリズム的情報理論の基礎となるノイズのないコード化、Chaitin マシンを用いたプログラムサイズの記述量 (複雑さ) などの基本的概念を調査、考察した上で、任意の Chaitin マシンや任意の semi-distribution に対して、アルゴリズム的コード化定理を満たすマシンのク

ラスを調査、考察する。さらに、アルゴリズム的コード化定理を満たしているすべてのChaitin マシンの特徴付けを行い、最適な (定数が 0 である)Chaitin マシンのクラスを示すことを目的とする。

2 ノイズのないコード化

ノイズのない通信路を通るメッセージをエラーがなくかつ可能な限り速く変換 (コード化、デコード化) するため方法として C.E.Shannon によって考案された prefix コードや prefix-free 集合の定義や性質を示す。その性質とは、prefix-code は唯一に復元可能であること、prefix-free 集合 S は Kraft の不等式 $\sum_{x\in S} 2^{-n} \le 1$ を満たすこと、さらに semi-distribution に対しても Shannon のノイズのないコード化定理を満たすことなどがある。

3 プログラムサイズの複雑さ (記述量)

情報の中身を測るために、Turing マシンの特殊なモデルであり、prefix-free な定義域を持っている Chaitin マシン M を用いる。対象 (文字列)x の記述量 $H_M(x)$ は、Chaitin マシンで x を計算するための最小プログラム p のビットサイズとして定義される。以下の定理は不変量定理と呼ばれ、AIT の起源になっている定理である。ある万能マシン U が存在して、任意のマシン M に対し、任意の対象 x に対し、以下の式を満たす定数 c が存在する。

$$H_U(x) \le H_M(x) + c.$$

不変量定理は対象 x が記述方法 U, M とは独立に決まることを意味している。そして、この U を固定して、簡単な上界を見積もることができる。他の重要な性質として、プログラムサイズの複雑さ H_M の計算不可能性、準計算可能性などが挙げられる。

AIT における重要な定理を二つ挙げる。

Kraft-Chaitin 定理: $\sum_i 2^{-n_i} \le 1$ を満たす枚挙可能集合 (r.e.set)$S = \{(x_i, n_i) \in \Sigma^* \times N \mid i \ge 0\}$ を与えると、$M(u_i) = x_i$ for all i を満たすマシン M を構成することができる。

この定理は Kraft の不等式を任意の枚挙可能集合に拡張したものである。もう一つは、Kraft-Chaitin Theorem を利用して導かれた、アルゴリズム的コード化定理である。

アルゴリズム的コード化定理; 任意の万能マシン U の記述量 $H_U(x)$ はマシンのアルゴリズム的確率 P_U に関して、漸近的に最適である (i.e. 定数で抑えられる);

$$H_U(x) \le -\log P_U(x) + (1 + c).$$

$$H_U(x) = \min\{|u| \mid U(u) = x\},\ P_U(x) = \sum_{U(u)=x} 2^{|u|}.$$

次のセクションでこの結果を一般的にする。

4 Characterization

与えられた semi-distribution P の状況のもとで、アルゴリズム的コード化定理を満たすマシン M を調査する。

基本的な結果：

> P を semi-distribution とし、$S \subset \Sigma^* \times N$ を枚挙可能集合とするとき、以下の二つの状況を満たすような定数 $c \geq 0$ が存在する。
> 状況：任意の $x \in \Sigma^*$ に対して
>
> *(1)* $\sum_{(x,n)\in S} 2^{-n} \leq P(x)$,
> *(2)* 任意の自然数 $n \in N$ に対して, もし $P(x) > 2^{-n}$, そのとき $(x, k) \in S$ かつ $k \leq n + c$ となる $k \in S$ が存在する。
>
> その時、任意の $x \in \Sigma^*$ に対し、以下の不等式を満たすマシン M が存在する。
>
> $$-\log P(x) \leq H_M(x) \leq (1 + c) - \log P(x).$$

この結果を利用し、準計算可能な semi-distribution、計算可能な semi-distribution を与えた時その条件が満たされているかを調べる。

P が下から準計算可能な semi-distribution の時、$c = 1$ で、アルゴリズム的コード化定理を満たすマシンを構成できる。P が計算可能な semi-distribution の時、$c = 0$ で、アルゴリズム的コード化定理を満たすマシンを構成できる。

最後に、アルゴリズム的コード化定理を満たすクラスのマシンは、以下の条件を満たすような $c \geq 0$ が存在するときである。
条件: 任意の自然数 n に対し, もし $P_M(x) > 2^{-n}$ ならば, そのとき $H_M(x) \leq n + c$ を満たすときである。。

定数 $c = 0$ でアルゴリズム的コード化定理を満たすクラスのマシンは、任意の異なる文字列 $u \neq u'$ に対し、$M(u) = M(u')$ ならば $|u| \neq |u'|$ を満たすときである。

参考文献

[1] Cristian.S.Calude, *Information and Randomness :An Algorithmic Perspective*, Springer-Verlag, 1994.

[2] Cristian.S.Calude, Hajime Ishihara and Takeshi Yamaguchi, *Minimal programs are almost optimal*, CDMTCS Reserch Report Series, CDMTCS-116, 1999, Spring.

[3] Greg.J.Chaitin, *Information, Randomness and Imcompleteness, Paper on Algorithmic Information Theory*, World Scientific, Singapore, 1990(2nd ed.,).